

Gravitational instabilities in a protosolar-like disc – II. Continuum emission and mass estimates

M. G. Evans,^{1★} J. D. Ilee,² T. W. Hartquist,¹ P. Caselli,³ L. Szűcs,³ S. J. D. Purser,¹
A. C. Boley,⁴ R. H. Durisen⁵ and J. M. C. Rawlings⁶

¹*School of Physics and Astronomy, University of Leeds, Leeds LS2 9LN, UK*

²*Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK*

³*Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse, D-85741 Garching bei München, Germany*

⁴*Department of Physics and Astronomy, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada*

⁵*Department of Astronomy, Indiana University, 727 East 3rd Street, Swain West 319, Bloomington, IN 47405, USA*

⁶*Department of Physics and Astronomy, University College London, London WC1E 6BT, UK*

Accepted 2017 May 31. Received 2017 May 25; in original form 2017 April 6

ABSTRACT

Gravitational instabilities (GIs) are most likely a fundamental process during the early stages of protoplanetary disc formation. Recently, there have been detections of spiral features in young, embedded objects that appear consistent with GI-driven structure. It is crucial to perform hydrodynamic and radiative transfer simulations of gravitationally unstable discs in order to assess the validity of GIs in such objects, and constrain optimal targets for future observations. We utilize the radiative transfer code *LIME* (Line modelling Engine) to produce continuum emission maps of a $0.17 M_{\odot}$ self-gravitating protosolar-like disc. We note the limitations of using *LIME* as is and explore methods to improve upon the default gridding. We use *CASA* to produce synthetic observations of 270 continuum emission maps generated across different frequencies, inclinations and dust opacities. We find that the spiral structure of our protosolar-like disc model is distinguishable across the majority of our parameter space after 1 h of observation, and is especially prominent at 230 GHz due to the favourable combination of angular resolution and sensitivity. Disc mass derived from the observations is sensitive to the assumed dust opacities and temperatures, and therefore can be underestimated by a factor of at least 30 at 850 GHz and 2.5 at 90 GHz. As a result, this effect could retrospectively validate GIs in discs previously thought not massive enough to be gravitationally unstable, which could have a significant impact on the understanding of the formation and evolution of protoplanetary discs.

Key words: protoplanetary discs – stars: pre-main-sequence – submillimetre: planetary systems – submillimetre: stars.

1 INTRODUCTION

Simulations have shown the surrounding protoplanetary disc can contain a mass comparable to that of the central protostar in newly formed systems (e.g. Machida & Matsumoto 2011). In this scenario, gravitational instabilities (GIs) can form in the disc and drive global spiral waves. As these spiral waves grow, they can produce shocks that heat the disc material locally (e.g. Harker & Desch 2002; Boley & Durisen 2008; Bae et al. 2014), which has a significant effect on the chemical evolution of some species as Ilee et al. (2011) and Evans et al. (2015, hereafter *Paper I*), have shown. As the dust emission within discs is dependent on the temperature, the spiral

shocks should produce a flux contrast between the arm and inter-arm regions. Therefore, continuum observations of GI-driven spiral structure in young, embedded systems would be of great importance as the formation and evolution mechanisms of protoplanetary discs are still uncertain.

There has been a continuous advancement in observational instruments, which has resulted in an unprecedented amount of protoplanetary disc images at millimetre wavelengths. However, because spatially resolving the youngest, most embedded sources (late Class 0/early Class I) at au scales (the expected size of GI-driven spirals) remains very challenging, most observations to date have focused on more evolved discs (Class II). Spiral features have been detected in Class II discs (e.g. Muto et al. 2012; Benisty et al. 2015), but Dong et al. (2015) raise several important points that challenge the validity of GI-driven spiral arms in more evolved objects. First, derived

*E-mail: py09mge@leeds.ac.uk

Table 1. Dust opacity values for different dust grain configurations across the frequency range used when deriving disc mass estimates.

| Dust grain configuration | Frequency (GHz) | | | | |
|--------------------------|---|-------|-------|-------|--------|
| | 90 | 230 | 300 | 430 | 850 |
| | κ_v (cm ² g ⁻¹) | | | | |
| ThinIceNoCoag | 0.112 | 0.509 | 0.782 | 1.488 | 5.827 |
| ThinIceCoag | 0.200 | 0.896 | 1.369 | 2.593 | 9.938 |
| NoIceCoag | 0.635 | 1.987 | 2.744 | 4.591 | 11.184 |
| D'Alessio et al. (2001) | 0.523 | 1.557 | 2.111 | 3.653 | 12.082 |

disc masses are not typically high enough to induce GIs. Secondly, the accretion rates in observed Class II objects do not appear consistent with the accretion rates in their simulations of gravitationally unstable discs. Thirdly, the observed features are probably beyond the critical radius where discs are believed to fragment. Dong et al. (2015) conclude by offering planet–disc interactions as a more credible origin for the spiral features, but it should be noted that there are counterarguments to each of the aforementioned points, and even planet–disc interactions may not explain observations to satisfaction (Richert et al. 2015).

Since the Atacama Large Millimetre/submillimetre Array (ALMA) became operational, the sensitivity of millimetre observations has increased, allowing us to peer deep into younger systems for the first time. As a result, Pérez et al. (2016) have detected spirals originating from the bulk of the disc surrounding Elias 2-27 that appear consistent with GIs (Meru et al. 2017; Tomida et al. 2017). Moreover, Tobin et al. (2016), have detected spiral features in a Class 0 object that also appear likely to have originated from GIs, albeit in a fairly complex multiple star system where the spirals may be produced by gravitational interaction between the nearby young stellar objects.

In light of these recent discoveries, it is currently a very exciting time in protoplanetary disc observations. However, even though spirals are beginning to be detected in real discs, it remains important to simulate observations, where all the properties are known a priori, in order to assess the validity of GIs in young discs. Furthermore, simulated observations will allow the constraint of the observational parameters necessary for modern technologies such as ALMA to unequivocally detect GI-driven spiral features. Several authors have reported that GI-driven spirals should be detectable with ALMA in discs with a range of masses and sizes (e.g. Cossins, Lodato & Testi 2010; Douglas et al. 2013; Dipierro et al. 2014), and we build upon this repository of results by synthesizing observations of the hydrodynamic simulation used in Paper I, which is a gravitationally unstable $0.17 M_\odot$ disc surrounding a $0.8 M_\odot$ protostar. This system may be analogous to our early Solar system, so the synthetic detection of spiral structure in this case would be of particular significance.

An important result derived from observations of protostar systems is the mass of the surrounding protoplanetary disc as it is a pivotal quantity in understanding disc formation and evolution. Estimates from observations of the nearby Taurus star-forming region have determined that Class II disc masses range between 0.0003 and $0.06 M_\odot$ (Andrews et al. 2013) and masses of Class 0 discs range between 0.001 and $0.6 M_J$ (see Greaves & Rice 2011, Table 1). These estimates are based on the assumption of vertically isothermal disc structure and optically thin dust emission, which is perhaps appropriate for lower mass discs, although even this is debatable as masses derived from accretion rates tend to be higher. However, for younger and more embedded discs these assumptions are more

likely to be inaccurate for a number of reasons. First, shocks driven by GIs heat and lift material from the mid-plane, hence the vertical structure is far from isothermal (see Boley & Durisen 2006; Evans et al. 2015). Secondly, due to the surrounding envelopes and high densities, Class 0/I discs are likely optically thick even at millimetre wavelengths (e.g. Miotello et al. 2014). As a result, the observed flux from very young systems would only be tracing a fraction of the actual disc mass. This has implications for existing and future studies as derived masses could be underestimated, which could retrospectively validate GIs in discs previously thought not to be gravitationally unstable.

In this paper, we take the radiative hydrodynamic model of a gravitationally unstable disc from Paper I and use it to investigate the observability of GI-driven spiral structure in dust continuum emission using ALMA. We establish the optimal parameters with which to perform the LIME (Line modelling Engine) radiative transfer, a crucial step for our next study on the observability of molecular species (Evans et al., in preparation), for which we calculate the chemical evolution in Paper I. We produce ALMA synthetic observations of our disc model at different frequencies and inclinations by implementing a range of dust opacities in the radiative transfer calculations. Our aim here is to provide an insight into how sensitive the interpretations of observations of gravitationally unstable discs can be to the assumed dust opacities, as the values for young, embedded objects are poorly constrained. We then compare observationally derived disc masses with the actual disc mass and compare assumed dust temperatures with the actual dust temperatures. We also provide a comparison between synthetic observations derived from LIME images to those derived from another existing radiative transfer code, RADMC-3D (Dullemond et al. 2012), because it is important to ensure there is consistency between different radiative transfer codes in the literature.

In Section 2, we briefly explain how LIME can be used to produce intensity maps and explore methods to improve upon the base LIME setup, which we refer to as ‘vanilla’ LIME. These include changing the weighting parameters; changing the number of grid points; restricting grid point positioning based on the optical depth; and the effect of averaging multiple runs. In Section 3, we produce synthetic observations of our disc model across a large parameter space and use the flux density to determine the disc mass at each combination of parameters. We then compare the observationally determined disc masses to our actual simulation mass to assess the validity of this commonly used observational method. Finally, in Section 4, we present our conclusions and discuss future research.

2 PRODUCING INTENSITY MAPS WITH LIME

2.1 Disc model

The disc model we use to produce continuum emission maps is a snapshot of the lower mass disc featured in Paper I, taken at the end of the simulation, $t \approx 2050$ yr, once the spiral features are fully developed and the disc is in a rough balance between heating and cooling processes, i.e. the disc is self-regulated and not fragmenting. Details of the hydrodynamic simulation can be found in Paper I, which we briefly describe here. The disc has a mass of $0.17 M_\odot$, spans approximately 50 au in radius, and the simulation was performed using 3D radiative hydrodynamics. The heating and cooling mechanisms consist of radiative energy losses, PdV work, viscous dissipation and irradiation by a central protostar. The regular grid consists of consists of 512, 512 and 64 cells with 0.25 au resolution in the x -, y - and z -directions, respectively.

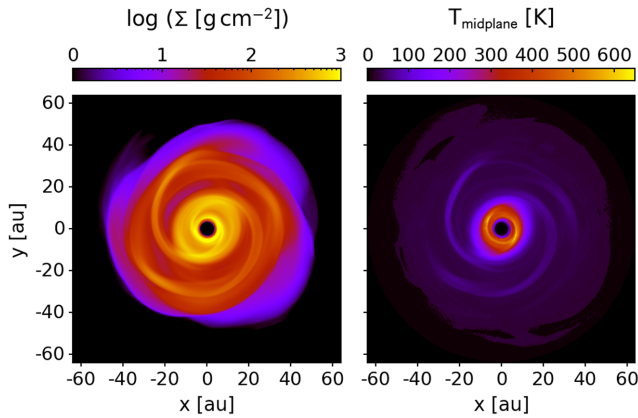


Figure 1. Nuclei column density map of the disc model (left-hand side) and mid-plane temperature map (right-hand side), showcasing the non-axisymmetric spiral structure in the disc model.

The dust model adopted in the hydrodynamic simulation assumes D’Alessio, Calvet & Hartmann (2001) opacities with a grain size distribution $n(a) = n_0 a^{-3.5}$ and $\alpha_{\text{max}} = 1$ mm to account for grain growth (D’Alessio et al. 2006). We use a gas-to-dust mass ratio of 100 and the dust grains are assumed to be thermally coupled and well mixed with the gas since we do not expect significant dust settling in such turbulent systems, and the computed Stokes number for mm grains is only greater than unity at the very edge of the disc.

We do not include an envelope in our radiative hydrodynamic simulation as we are only focusing on the detectability of GI-driven spiral structure. In reality, a surrounding envelope may shield the disc from external radiation that can otherwise diminish the spiral structure (e.g. Cai et al. 2008; Kratter & Murray-Clay 2011), but here we are assuming our disc is already well shielded, i.e. heating from the interstellar radiation field is neglected. Furthermore, infall from the envelope on to the disc can significantly affect the spiral structure (e.g. Harsono, Alexander & Levin 2011), but as the consequence appears to be an enhancement in the contrast between arm and inter-arm regions, neglecting this interaction will only be significant for our conclusions if we do not detect GI-driven spirals.

We do, however, incorporate an envelope into the radiative transfer step of our modelling when producing continuum emission maps. In order to appropriately describe the environment that a young protoplanetary disc is embedded within, we take this envelope to be a $10 M_{\odot}$ contracting Bonnor–Ebert sphere (see Keto, Rawlings & Caselli 2014). Moreover, whilst radiative transfer codes often self-consistently calculate the dust temperature, this only considers radiative heating. Instead, we utilize the dust temperatures from the radiative hydrodynamics code, which combines flux-limited diffusion with raytracing in the vertical direction (see Boley et al. 2007), in order to account for the viscous and shock heating in our model. As a result we only need to perform raytracing on our disc model to obtain continuum emission maps.

The nuclei column density map and mid-plane temperature of the disc snapshot, we utilize are shown in Fig. 1 to enable visual comparison to the continuum emission images presented in this paper.

2.2 ‘Vanilla’ LIME

We use the LIME (Brinch & Hogerheijde 2010), which calculates continuum emission and line intensities from a weighted Monte Carlo sampling of an input 3D model. Here we provide a brief

description of how LIME operates, but we refer the reader to Brinch & Hogerheijde (2010) and Douglas et al. (2013, section 2.3) for more detailed descriptions.

The input model contains density and temperature information, and also abundance and velocity information if considering line emission. LIME then selects an x -, y -, z -position randomly, and if a selection criterion is met, which is dependent on the density by default, this point is added to the grid. Once the grid is constructed, LIME calculates the appropriate parameter values at each point; the method used to achieve this, such as nearest point interpolation, linear interpolation or something more advanced, is defined by the user in the model file. The constructed grid is then smoothed via Lloyd’s algorithm (Lloyd 1982; Springel 2010) in order to ensure the distance between points is comparable to the local separation expectation value whilst still maintaining the underlying model structure. The grid is then Delaunay triangulated and if non-local thermodynamic equilibrium (non-LTE) line emission is being considered, photons are propagated along the Delaunay lines and the radiative transfer equation is solved at each grid point via an iterative process. Finally, once convergence is reached, or if only continuum emission is being considered, LIME ray-traces lines of sight through the model in order to produce intensity maps.

LIME is in continual development, has been hosted on GitHub since v1.3¹, and the available documentation is rapidly improving. For now, LIME offers the user the ability to specify a whole array of parameters that affect the resultant intensity maps, which can be complex for first time users. In order to highlight potential pitfalls, we ran LIME v1.6 as is, which we refer to as ‘vanilla’ LIME, for our disc model using 2.5×10^4 grid points; we use this number of grid points initially as the documentation recommends ‘between a few thousands up to about one hundred thousand’ and it is also the number of grid points used in Douglas et al. (2013).

The leftmost panel of Fig. 3 shows the 300 GHz continuum emission map for our disc model (see Paper I) when using vanilla LIME, and comparing to Fig. 1 it is immediately apparent that the non-axisymmetric structure is very poorly resolved. This is because the default weighting routine sets a normalization density at the model inner boundary, which in our model is within the inner hole. This means we are essentially omitting any density weighting when using vanilla LIME. Note, we adopt a frequency of 300 GHz for the rest of Section 2.

2.3 Weighting of grid points

In order to procure a more accurate flux image, we must amend the density weighting function to be appropriate for our model. LIME randomly distributes points during the grid building routine and only selects the point if it passes a criterion

$$a < \left(\frac{n}{n_0} \right)^w, \quad (1)$$

where a is a random number generated between 0 and 1 each time a point is selected, n is the density of the grid point, n_0 is a normalization density and w is an exponent. The normalization density was fixed to the inner boundary density in LIME v1.5 and earlier, but with v1.6 the user now should set the parameter n_0^w in the model file. If the user does not set this parameter, however, then the old default is restored. This can be a poor selection criterion if, for example, the model has an inner hole or if the user would like to

¹ <https://github.com/lime-rt/lime>

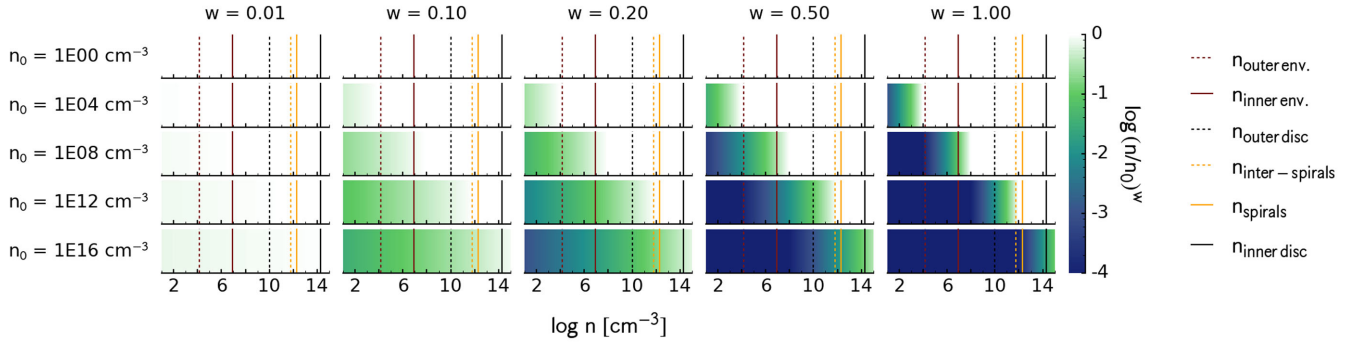


Figure 2. Probability of grid point selection as the density weighting parameters, n_0 and w , are varied. The vertical lines correspond to approximate number densities within various features of the disc model, as labelled.

focus on features with a lower density than the central density. We demonstrate this effect in Fig. 2 that shows the probability of point selection, $\log(n/n_0)^w$, as n_0 and w are varied, with features of our disc model overlaid. As can be seen, if the reference density, n_0 , is increased then the likelihood of point selection at low densities becomes increasingly small, and if the exponent, w , is increased then the range where points are likely to be selected contracts. For our purposes, we want a high probability of point selection at typical spiral arm densities, indicated by the solid orange vertical lines, but also with a significant percentage of points extending to the outer disc regions, with an approximate density denoted by the dotted dark grey lines. Therefore, from Fig. 2 we determine that $n_0 = 1 \times 10^{12} \text{ cm}^{-3}$ and $w = 0.5$ are optimal choices for our disc model. Note that this choice of parameters results in a small fraction of points being positioned within the envelope. This is not an issue, however, because the maximum envelope density is lower than the bulk of the disc model. As a result, the mean free path within our simplistic envelope model is ubiquitously larger than the envelope size and hence we only need a small fraction of points to treat the emission accurately. In reality, the innermost part of the envelope around embedded protoplanetary discs may be very dense, which could affect our conclusions in the outer disc regions.

We ran vanilla LIME for our disc model, which equates to using $n_0 = 100 \text{ cm}^{-3}$ and $w = 0.2$. We then ran LIME using $n_0 = 2 \times 10^{14} \text{ cm}^{-3}$ and $w = 0.2$ to simulate the intention of the default reference density value assignment. Finally, we ran LIME using $n_0 = 1 \times 10^{12} \text{ cm}^{-3}$ and $w = 0.5$, which, as aforementioned, we determined from Fig. 2. Fig. 3 shows the flux image outputs of these runs and as can be seen, if n_0 is set much lower than the typical disc density then essentially no weighting is applied, apart from excluding the inner hole. In this case, there is poor sampling of the densest features and a large percentage of points are positioned in regions of low density. If, on the other hand, we set n_0 to the highest density in our model, as shown in the middle panel, then density weighting is applied and the disc spirals and inner region are sampled much more thoroughly. However, if we compare this result to the result in the right panel, where n_0 is set at the approximate spiral density, we can see that using too high a reference density undersamples the spiral features comparatively. This is because, as Fig. 2 shows, the probability of point selection at the density of the spiral arms is reduced as n_0 increases. Indeed the panel on the right-hand side, with $n_0 = 1 \times 10^{12} \text{ cm}^{-3}$ and $w = 0.5$, showcases a type of ‘Goldilocks’ regime for our disc model, in which we sample the disc spirals in an optimal manner. Note, though, that the trade-off is a smaller percentage of points at the highest densities, which in our model corresponds to the innermost disc, as indicated by the inset histogram in the right

panel of Fig. 3. Therefore, the optimal choice of n_0 is dependent on the features of interest in the model. For our purposes, we use $n_0 = 1 \times 10^{12} \text{ cm}^{-3}$ and $w = 0.5$ and refer to this setup as our standard sampling hereafter.

2.4 Number of grid points

Due to the Monte Carlo nature of the LIME gridding, two runs using identical inputs will likely produce differing images as the grid points will be positioned differently. This effect obviously decreases as the number of grid points is increased, so in order to circumvent this issue, the user can simply use a large number of grid points. For continuum images this is not a particularly significant issue because the computational cost is relatively low. However, for line images, because each molecular energy level population must be computed for each grid point within each iteration, the computational cost increases dramatically with the number of grid points. Therefore, even though this paper focuses on continuum images, deducing the minimum number of grid points we can use and still obtain accurate images in the continuum is a worthwhile endeavor as the results could be applicable to line images.

To achieve this we adopt the weighting parameters discussed in Section 2.3 and produce continuum images with an increasing number of total grid points until the residual between two runs using identical inputs met a sufficient level. We set this level to 5 per cent because this is smaller than the errors expected in observations. We find that the majority of the residual flux is smaller than our threshold level when using 2×10^6 points, as Fig. 4 shows. Note that the large residual differences seen towards the edges of Fig. 4 are where the model transitions from the disc to the envelope and the abundance of points drops off sharply. As we are focusing on the spiral structures we omit this region in future analyses.

By comparing the images produced using fewer number of grid points with this ‘canonical’ result, Fig. 5 shows that increasing the number of grid points increases the accuracy of the residual images. This is obviously a trivial result, but more importantly, Fig. 5 indicates that a sufficiently accurate image, which we have defined prior as the majority of the disc having a residual flux lower than 5 per cent, can be produced using 2×10^5 grid points. This is a factor of 10 lower than our ‘canonical’ result, and hence affords us a significant reduction in computational cost when producing continuum images. Not only is this a benefit in reducing computational time, but it also reduces memory consumption which allows more instances of LIME to be run simultaneously, which is utilized in Section 2.6.

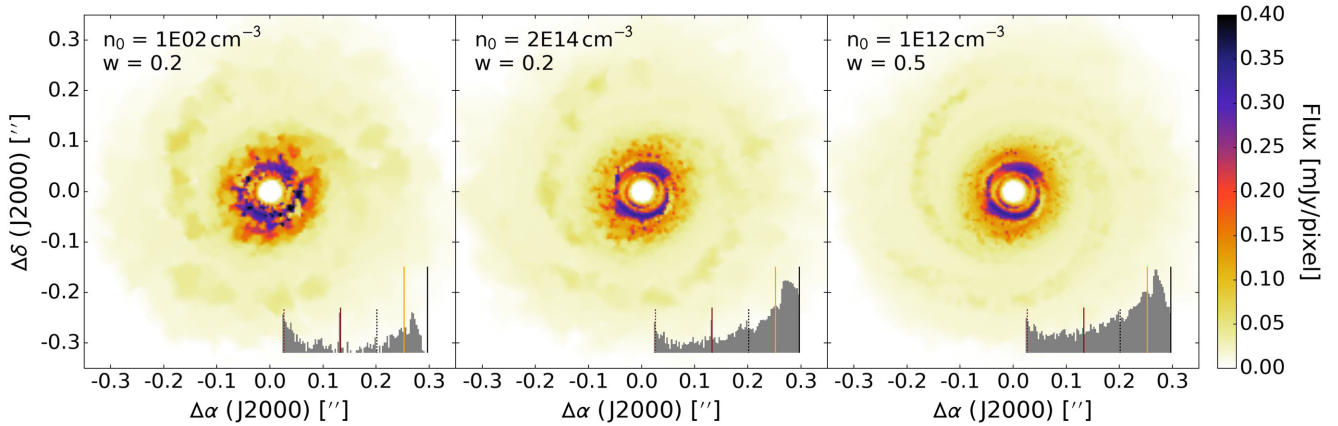


Figure 3. Comparison between 300 GHz flux images produced by LIME using varied n_0 and w parameters, demonstrating that LIME images are sensitive to the weighting imposed. The inset figures show the distribution of number densities of the grid points, confined within the disc region, that produce the corresponding image, with the vertical lines corresponding to the key in Fig. 2.

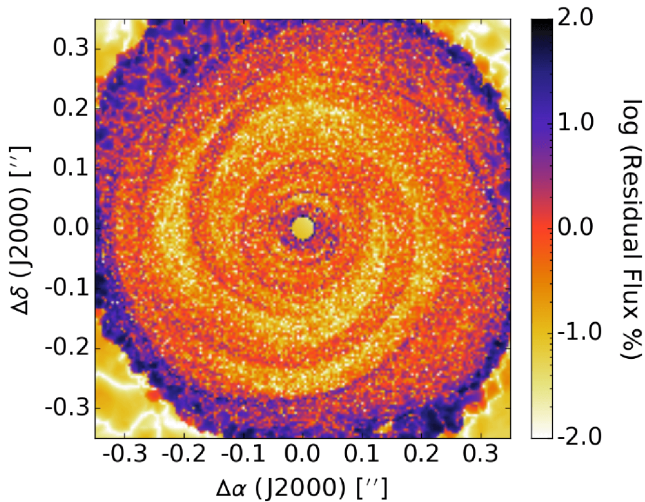


Figure 4. Residual flux between two LIME runs using identical inputs with 2×10^6 grid points and standard sampling. We define this residual as $|(f_1 - f_2)/(f_1 + f_2)|$, where f denotes flux per pixel.

Although the results demonstrated in Fig. 5 are expected, it is important to understand precisely why an increase in grid points affords an improvement in the image accuracy. The emission of photons is governed by the mean free path, which is a statistical average of the distance a photon will travel before interacting. If the grid point separation is larger than the local mean free path, then important photon interactions will be omitted. Hence the resultant flux images will be inaccurate.

The mean free path is given by $l = 1/\alpha_v$, where α_v is the absorption coefficient at a particular frequency. For dust-dominated continuum emission $\alpha_v = \kappa_v \rho_d$, where κ_v is the dust opacity at a specific frequency and ρ_d is the dust mass density. We know the κ_v value as it is contained within a dust opacity file read by LIME, and we also have the gas mass density information within our hydrodynamic simulation, which we convert to dust mass density using the commonly adopted value of 100 for the gas-dust mass ratio. Using these quantities, we calculate mean free path values throughout our disc model in an analytical fashion. Next, we straightforwardly obtain the grid point separation information as LIME v1.6 has the ability

to output the grid point positions and nearest neighbour distances. Finally, we interpolate the nearest neighbour distances to a coordinate grid matching our mean free path data cube and compare these length-scales.

We plot the comparison between the mean free path and grid point separation across the x - y plane at $z = 0$ for each of the runs shown in Fig. 5. The results are shown in Fig. 6 and demonstrate that by increasing the number of grid points, the disparity between these two length-scales across the majority of the disc is reduced. However, even when using 1×10^6 points there is still an appreciable discrepancy between the mean free path and grid point separation within the spiral arms and inner disc regions. This is not surprising because these are the densest regions of our disc model and hence have the smallest local mean free path values. Fortunately, this is not a particularly large issue because the contribution to intensity in these densest, most optically thick regions is small. In fact, we can account for this phenomenon and optimize our sampling routine further.

2.5 Optical depth surface

In optically thick regions, such as those expected in embedded protoplanetary discs, the contribution to observed intensity is low because $I_{\nu,0} \propto e^{-\tau}$. Therefore, we designate a threshold of $0.05I_{\nu,0}$, below which point we assume the emission becomes negligible as this is less than typical observational errors concerning embedded objects. This means we should only need to place grid points up to a surface where the optical depth reaches $\tau = -\ln(0.05) = 3$ in order to produce an accurate image. We refer to this surface as the optical depth surface hereafter. For reference, the photosphere of a star is defined at $\tau = 2/3$, hence we are likely still oversampling the densest regions of our disc model by adopting $\tau = 3$ for our optical depth surface. However, determining the minimum optical depth one can use in grid point restriction and still obtain accurate images is dependent on dust opacity and frequency, which requires a dedicated study and as such is beyond the scope of this work.

LIME is capable of outputting several different image types, such as brightness temperature, flux and optical depth along the line of sight, which the user can specify; we adapt the way LIME outputs images so that the user can now choose to have multiple image types output for one set of image parameters, which has been incorporated into LIME

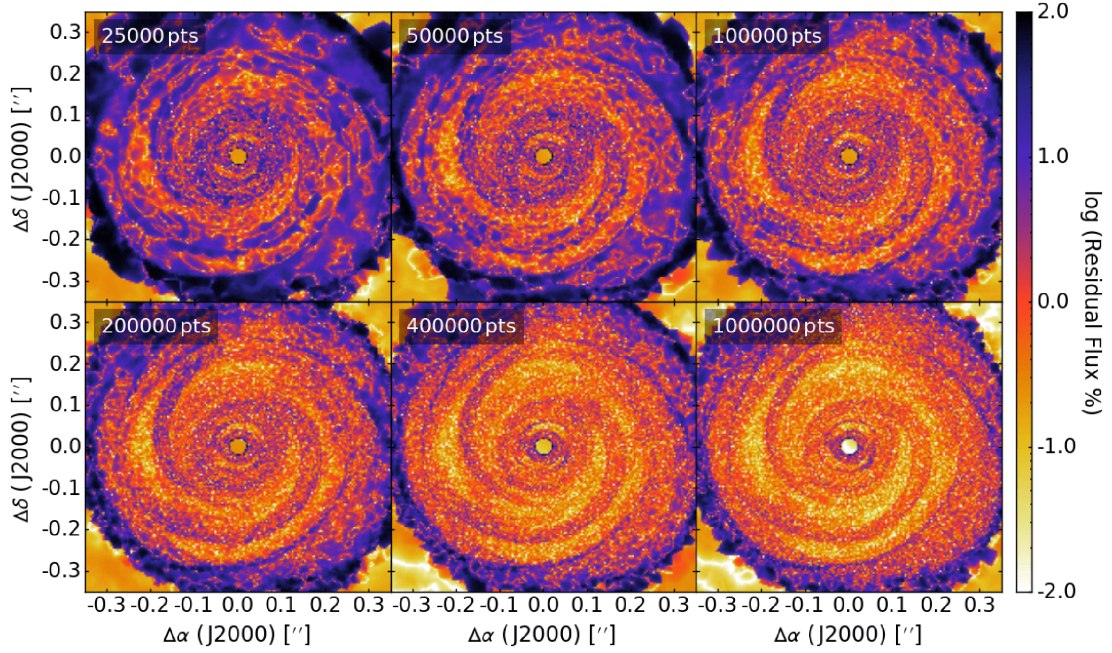


Figure 5. Residual flux between standard LIME runs using the indicated number of grid points and 2×10^6 grid points, with standard sampling. We define this residual as $|(f_1 - f_2)/(f_1 + f_2)|$, where f denotes flux per pixel.

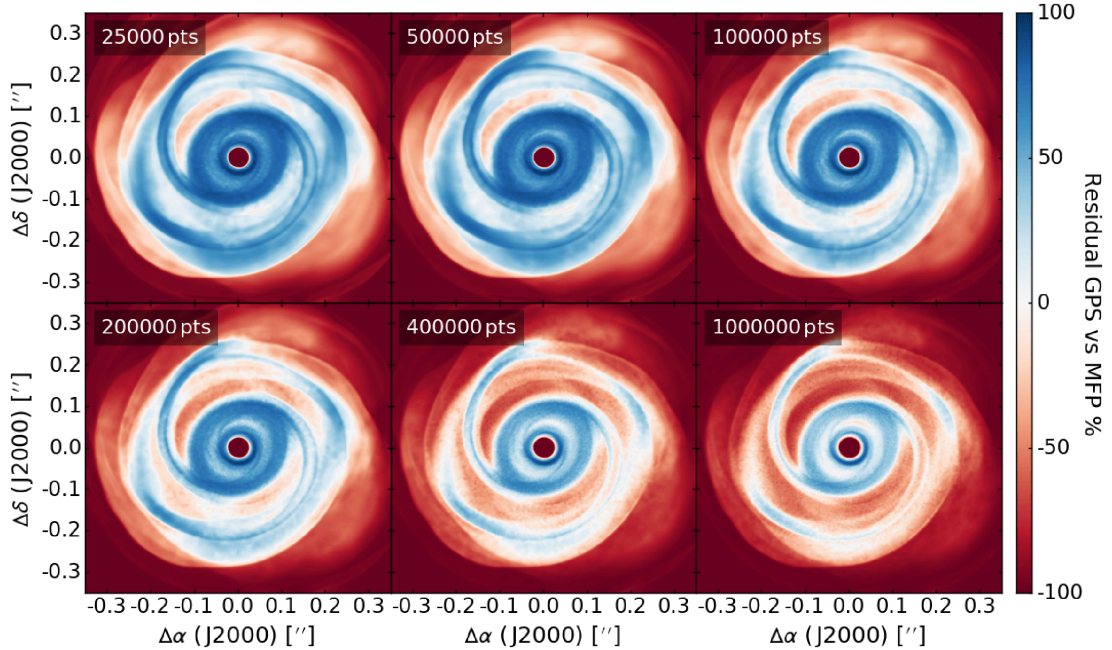


Figure 6. Residual of grid point separation, s , and mean free path, l , at the disc mid-plane for LIME runs using the indicated number of grid points and standard sampling. We define this residual as $(s-l)/(s+l)$.

v1.7.2. In order to implement our optical depth surface method,² we add a new image output to LIME that allows us to determine where in our disc model this optical depth surface lies. This is achieved via an addition to the ray-tracing routine that records the z position for each propagated ray once the optical depth has surpassed our

specified value. The optical depth is calculated via $\tau = \kappa_v \rho_d ds$, where κ is the dust opacity at the specified frequency (300 GHz), ρ_d is the dust density and ds is the distance between Voronoi cells after each iteration of the ray propagation. As the optical depth can jump significantly beyond our threshold value from one Voronoi cell to the next, we implement linear interpolation. We pass this value into an additional ray *struct* parameter and then average over the number of rays that pass through each pixel in the same manner

² github.com/lolmevans/lime/tree/optical-depth-surface

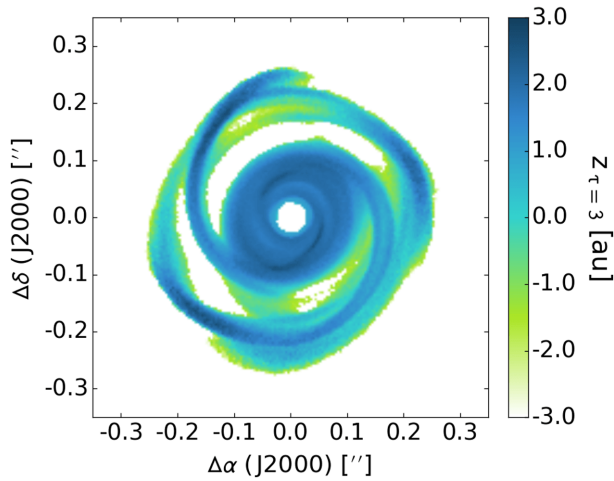


Figure 7. Surface across the disc model where $\tau = 3.0$ when viewed from above. Below this surface the contribution to intensity becomes insignificant.

as the existing flux and optical depth images. The resultant image displays the height above or below the mid-plane of our desired optical depth surface across the entire disc model.

Fig. 7 shows the $\tau = 3$ surface for our disc model when viewed from above. The optical depth surface is significantly above the mid-plane within the spiral arms due to the large column densities within the spiral arms, and because the interaction between disc material and the spiral shocks is similar to a hydraulic jump (Boley & Durisen 2006). Hence, material in the spiral arms is elongated vertically. Between the spiral arms, the optical depth surface lies below the mid-plane, which means we are peering through significantly more disc material in these regions.

As only about 5 per cent of the photons emitted at the $\tau = 3$ surface will escape, we can see with the use of Fig. 7 that the omission of grid points below $z \approx 3.0$ au in the spirals, below $z \approx 2.0$ au across the inner disc and below $z \approx -1.0$ au in the inter-arm regions is reasonable. Rather than use these approximate values, however, we amend the LIME gridding function to read in the disc optical depth surface and use this to restrict the grid point positions; if a point is selected at $\tau > 3$, it is translated to a region of $\tau < 3$ and re-evaluated. Effectively, we move the focus of points from

around the mid-plane to regions of $\tau \leq 3$ where the contribution to intensity is most significant.

We use this technique to reproduce the LIME image consisting of 2×10^5 grid points, calculate the residual from 2×10^6 points, and compare this to the bottom left-hand panel of Fig. 5. The results are shown in the left-hand and middle panels of Fig. 8 and as can be seen, limiting the grid points to above the optical depth surface affords an improvement in the image overall. In fact, when comparing to Fig. 5 we can see that the effect is similar to doubling the number of grid points, but without the large increase in computational time; implementing our method has no significant effect on the runtime but doubling the number of grid points increases the runtime by 400 per cent. It should be noted that in order to implement our optical depth surface routine, we need to run LIME initially to generate the optical depth lookup table. Despite this, however, our method is still more efficient than producing a continuum image with double the number of grid points. Therefore, this result is a proof of concept that positioning points more intelligently can result in more efficient radiative transfer calculations using LIME.

Although there is a general improvement visible between the left-hand and middle panels of Fig. 8 there are also some regions where the residual has deteriorated. This is particularly evident in the prominent dark feature above the inner hole. However, the residual difference in this region is approximately 5 per cent, which is within the expected threshold of omitted emission when using our $\tau = 3$ surface to constrain point positions. Furthermore, this residual difference equates to an absolute difference of 2 K and is located within a weakly emitting region. As a result, this rather visually striking difference is entirely insignificant, but we offer a possible explanation for completeness: there is a particularly sharp density gradient in this region and hence our optical depth method is not treating this area properly. This then implies that the density weighting employed by LIME is not entirely accurate for our method and a better alternative could be a weight dependence on the density gradient. This, however, is beyond the scope of this paper and is left as a future task, especially considering the small residual difference present here.

The reason this method produces a more accurate image is entirely consistent with our results from Section 2.4, where we increase the total number of grid points, which reduces the grid point separation throughout the entire disc. In this section, however, we are increasing the number of grid points in the emitting regions using selective positioning. This therefore reduces the disparity between

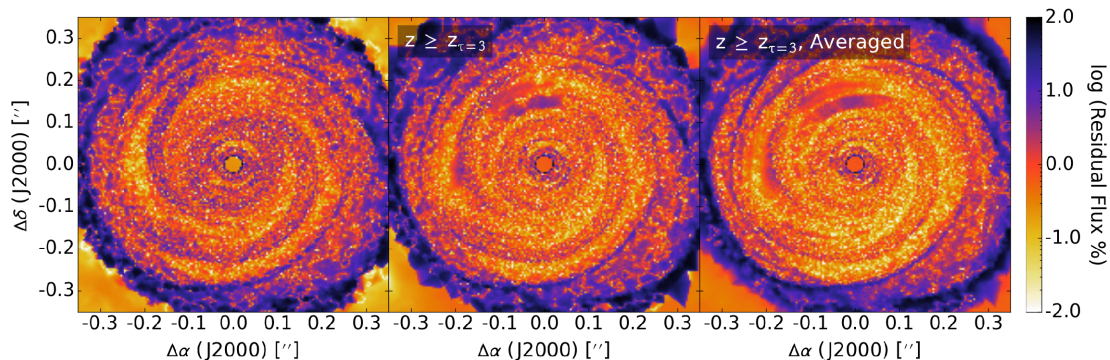


Figure 8. Residual flux between LIME runs using 2×10^5 grid points and different sampling methods, f_1 , and 2×10^6 grid points with standard sampling, f_2 . We define this residual as $|(f_1 - f_2)/(f_1 + f_2)|$, where f denotes flux per pixel. The panel on the left is the residual result for our standard sampling seen in Figure 5. The middle panel shows the residual result when using an optical depth surface to constrain the positioning of grid points, which we refer to as our optimal sampling. The right-hand panel shows the residual result for an average of eight optimally sampled LIME runs using identical input parameters.

grid point separation and mean free path within the emitting regions only, i.e. at $z \geq z_{\tau=3}$.

2.6 Averaging runs

Multiple instances of LIME can be run simultaneously, limited only by the number of processor cores and memory available. The averaging of ten runs using identical inputs was performed in Douglas et al. (2013) (see their fig. 6), which reduced image artefacts and smoothed their flux emission successfully. We adopt the same technique but use our optimal image setup ($n_0 = 1 \times 10^{12} \text{ cm}^{-3}$, $w = 0.5$, 2×10^5 points, $z \geq z_{\tau=3}$) and only average eight runs due to computational restraints.

The comparison between a single run and the averaging of multiple runs using identical inputs is shown in the middle and right-hand panels of Fig. 8, and as can be seen, averaging affords us an improvement to the output image. Due to the Monte Carlo nature of the LIME gridding, averaging multiple identical runs effectively increases the point coverage. This is essentially identical in effect to increasing the number of grid points, which, as shown in Section 2.4, improves the image due to the disparity reduction between the grid point separation and mean free path. It is worth noting that the improvements seen between the middle and right-hand panels of Fig. 8 are small in magnitude and likely much less than observational errors. However, because multiple instances of LIME can be run simultaneously, which affords a significant reduction in computation time compared to increasing the number of grid points, there seems to be no strong argument against using the average of the outputted images for further analysis. Furthermore, although we are only focusing on continuum images, which have an approximately linear increase in computation time with increased grid points, for line images this relationship is non-linear due to the iterative calculation of level populations. As a result, the benefit of averaging multiple runs using identical inputs should be emphasized when producing line emission images (Evans et al., in preparation).

3 SYNTHETIC OBSERVATIONS

Tobin et al. (2016) detected spiral structure in a very young disc that hosts multiple protostars, which appears consistent with a gravitationally unstable disc that recently underwent fragmentation. Moreover, Pérez et al. (2016) detected spiral structure in a single-protostar circumstellar disc that does appear consistent with GI-driven spiral arms (Tomida et al. 2017). Although the object featured in Tobin et al. (2016) is not consistent with our model, and the simulated model of Tomida et al. (2017) does not match some features of the observation in Pérez et al. (2016), these results offer the first potential confirmation of GI-driven spiral structure in protoplanetary discs, and suggest more will follow in the near future. Assessing the validity of GIs in such objects is crucial and therefore we produce synthetic continuum emission maps of our GI-driven disc model in order to offer potential insight into this conundrum.

We produce 270 LIME images of our disc model across a 3D parameter space. We use 4×10^5 grid points and choose not to implement the optical depth surface method described in Section 2.5 as we would need to produce a $\tau = 3$ surface, and then run LIME again whilst implementing this surface. As we are only considering continuum emission using pre-computed dust temperatures, we only need to perform raytracing. Hence, the increase in image accuracy when adopting our optical depth surface model is offset by the significant increase in computational time from running LIME

twice for each combination of dust opacity and frequency. Moreover, by comparing Figs 5 and 8, we can see that a single run of 4×10^5 points is comparable to our optimal image (produced using the average of eight runs implementing our optical depth surface method). Therefore, for this large parameter space study, we opt to use double the number of grid points without implementing our optical depth surface method in the interest of efficiency, whilst still producing accurate images for analysis. However, we emphasize that the purpose of Sections 2.5 and 2.6 is to demonstrate that the gridding routine *can* be optimized for a particular model, which will be much more prevalent when producing line images as the computational time no longer scales linearly with the number of points due to the necessity in calculating level populations. We will investigate this in a forthcoming publication.

The parameter space we consider consists of varying frequencies, inclinations and dust opacities. We use five different frequencies commonly implemented in observations of protoplanetary discs, 90, 230, 300, 430 and 850 GHz, equivalent to 3.33, 1.30, 1.00, 0.70 and 0.35 mm, respectively, and the inclinations we use range from face-on (0°) to edge-on (90°) in 5° increments. We use three dust grain properties: coagulated grains with ice mantles (ThinIceCoag); coagulated grains without thin ice mantles (NoIceCoag); and non-coagulated grains with thin ice mantles (ThinIceNoCoag), all taken from Ossenkopf & Henning (1994), where ‘coagulated’ refers to a coagulation after 10^5 yr for a gas density of 10^6 cm^{-3} .

3.1 Dust opacities

The values of the dust opacities for the different dust models are shown in Table 1 at the frequencies we use, alongside the D’Alessio et al. (2001) opacities used in the radiative hydrodynamic simulation. As can be seen, the ‘NoIceCoag’ dust model is consistent (within a factor of 1.3) of the D’Alessio et al. (2001) opacities across the observational frequency range we use (90–850 GHz), whereas the ‘ThinIceCoag’ and ‘ThinIceNoCoag’ dust models diverge substantially at low frequencies. For this reason, we acknowledge that synthetic observations produced using ‘ThinIceCoag’ and ‘ThinIceNoCoag’ dust opacities will not be entirely self-consistent with the disc simulation. However, we present these results as an exploration of the effects dust opacity can have on the observability of spiral structure. In truth, the evolution of a gravitationally unstable disc is affected by the heating and cooling processes (see Kratter & Lodato 2016), so repeating the radiative hydrodynamic simulation with, for example, the ‘NoIceCoag’ dust model would most likely change the spiral structure in our disc model. In this case, as Fig. 10 shows, the ‘NoIceCoag’ opacities in Table 1 are higher than the D’Alessio et al. (2001) opacities at micron wavelengths, which would result in an increased cooling time. This suggests that the contrast between arm and inter-arm regions would be less pronounced (Cossins, Lodato & Clarke 2009), hence we would expect the detection of spiral structure to become more difficult. However, whilst self-consistently exploring the effects of dust opacity on the observability of spiral structure in a 3D radiative hydrodynamic simulation is an important topic of study, it is an extremely computationally expensive process and is beyond the scope of this paper.

Whilst the dust models in Table 1 do not follow a simple power law $\kappa_\nu \propto \nu^\beta$ across our frequency range, we can extract β values at specific frequencies to allow comparisons to values derived in the literature. For instance, $\beta_{230 \text{ GHz}} = 2.02, 1.64$ and 1.10 for the ‘ThinIceNoCoag’, ‘ThinIceCoag’ and ‘NoIceCoag’ dust models, respectively. Miotello et al. (2014) find $\beta \approx 0.5$ – 1.0 in Class I objects which may suggest a higher level of grain growth than we consider.

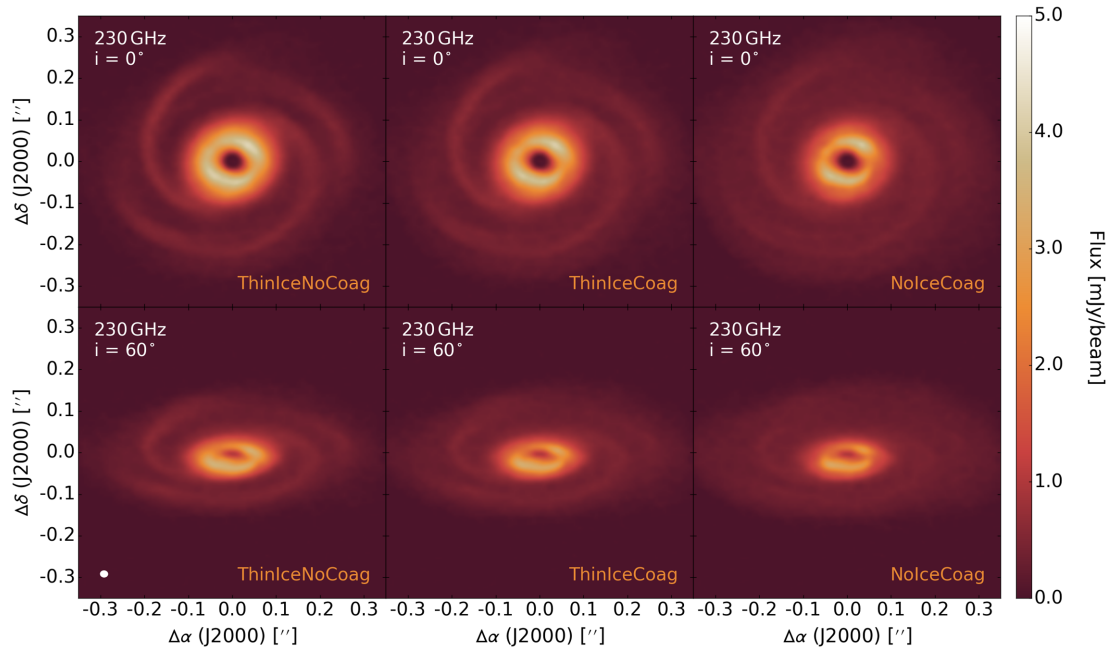


Figure 9. Continuum emission images of our disc model at 230 GHz for three different dust opacity models and two different inclinations, synthesized using the ALMA Cycle 5 antenna configuration. The white ellipse in the lower left-hand side indicates the size of the beam, which is 0.019×0.017 arcsec² and is constant across all panels. The noise is approximately $15 \mu\text{Jy beam}^{-1}$ across the parameter space.

However, the relative contribution of envelope and disc to this value is not clear, and the change in the dust grain population between the envelope and disc is not yet well characterized, especially in gravitationally unstable systems.

3.2 Detecting spiral structure

In order to investigate whether spiral structure can be distinguished across our parameter space, we synthesize observations of our LIME images using CASA (v4.5.0; McMullin et al. 2007). We assume our model is located at a distance of 145 pc in order to examine the observability of self-gravitating discs in the nearest star-forming regions (e.g. ρ Ophiuchi). We vary the inclination of our disc from 0° to 90° to cover the range of orientations seen in observations of real systems such as IRAS 16243–2422, which is a protostellar binary with two sources, A and B separated by approximately 600 au, with near face-on and edge-on disc inclinations, respectively. We use two different antenna setups in order to simulate the capabilities of ALMA Cycle 5 and a fully extended, maximally operational ALMA. For ALMA Cycle 5, we use the antenna configuration that is available and affords us the best angular resolution at each frequency, i.e. 90 GHz: C43-10, 230 GHz: C43-10, 300 GHz: C43-8, 430 GHz: C43-7, 850 GHz: C43-7. We simulate observing for 60 min using the SIMOBSERVE routine in CASA, considering thermal noise with an ambient temperature of 270 K and a precipitable water vapour given when selecting the ‘automatic’ option in the ALMA Sensitivity Calculator,³ which uses the specified frequency to calculate appropriate values, i.e. 90 GHz: 5.186 mm, 230 GHz: 1.796 mm, 300 GHz: 1.796 mm, 430 GHz: 0.913 mm, 850 GHz: 0.658 mm. This observation time is expected to be sufficient for detecting the dust continuum emission in the spiral features for the majority of our parameter space. Finally, we use the multiscale

CLEAN algorithm for deconvolution of the synthetic visibility data, in order to effectively recover the extended disc and spiral arm structure.

Here, we present only the most detailed image attained when observing our disc model with ALMA Cycle 5, as seen in Fig. 9, and compile the results at the remaining frequencies in Appendix A. The results when using the fully extended, maximally operational antenna configuration are also not displayed here because our overall conclusions are the same, but the images can be found in Appendix B. The most detail is recovered at 230 GHz because there is an optimal combination of angular resolution and sensitivity affording a signal-to-noise ratio (S/N) of approximately 30 within the resolved spirals; we calculate the noise level by taking the rms of regions within the residual maps free from disc emission or sidelobes. Our observations show that spiral structure is sensitive to the assumed dust opacities, as the contrast between the arm and inter-arm regions is different between the panels, which is important because the state of grain growth is relatively unknown at such early stages of disc evolution. However, as the spiral structure is clearly visible within all panels of Fig. 9, relatively short observations should be capable of detecting spiral structure in young, gravitationally unstable discs across a wide range of inclinations and possible dust grain opacities.

We note here that it is important to cross-check the results of radiative transfer codes in order to ensure there is a sufficient degree of consistency. In order to assess the validity of our results obtained with LIME, we also perform a comparison with synthetic observations produced using RADMC-3D with regular gridding, which we detail in Appendix C. We find that there is a general level of consistency between the flux emitted from the spiral arms, though RADMC-3D produces a stronger contrast between arm and inter-arm regions that results in an easier detection of spiral structure within our disc model. Nevertheless, this only serves to strengthen our conclusions that ALMA should be able to detect spiral structure in a gravitationally unstable disc.

³ <https://almascience.nrao.edu/proposing/observing-tool>

The synthetic observations we present in this paper, including those presented in Appendices A, B and C, are encouraging for future continuum observations as they show that ALMA can detect non-axisymmetric structure in young, embedded systems. As aforementioned, this has already been proven with recent observations, however, we can draw further conclusions from our synthetic results in particular. First, unlike real observations to date, we know unequivocally that the spiral features are driven by GIs in our disc model and have successfully demonstrated ALMA can resolve such GI-driven spirals. Secondly, we observe our disc for only 60 min and can already resolve the spiral structure across most of our parameter space at low frequencies. If a longer observational time was adopted, then the spiral features would also become more prominent at higher frequencies, allowing the detection of spirals in a young, gravitationally unstable disc across multiple ALMA bands. This would allow determinations of the spectral index and give insight into the grain size distribution (e.g. Pérez et al. 2012), shedding light on processes such as grain growth and dust trapping (e.g. Dipierro et al. 2015) in young, embedded discs. Finally, our model is of a protosolar-like disc, which means ALMA should be able to resolve spiral features in a relatively low-mass disc that is possibly analogous to our early Solar system.

3.2.1 Comparison with other studies

The synthetic observations of our disc model show that ALMA can detect non-axisymmetric structure in young, embedded systems across a range of frequencies, inclinations and dust opacities. This agrees with Dipierro et al. (2014), who use a 3D smoothed particle hydrodynamics (SPH) radiative hydrodynamic code (Lodato & Rice 2004) and find that for a gravitationally unstable disc located in the TaurusAuriga and Ophiucus star-forming regions (≈ 140 pc), the spiral structure is readily detectable by ALMA across a range of frequencies, inclinations and disc-to-star mass ratios. However, it should be noted that their model uses a cooling prescription that does not take into account the irradiation emitted by the central star, which implies that the amplitudes of their spiral perturbations are overestimated. Countering this effect somewhat, though, is the fact that Dipierro et al. (2014) assume more grain growth has occurred, and hence use larger dust opacities than we consider, which we have shown reduces the detectability of spiral structure. Dong et al. (2015) use a 3D SPH simulation augmented by a hybrid radiative transfer in order to model both global cooling and radiative transfer (Forgan et al. 2009), which is comparable to the technique we implement. They find that spiral arms should be detectable in near-infrared scattered light observations of systems with $q \geq 0.25$, which our disc model is near the threshold of, and also find that their high- q models closely resemble real objects, though the validity of GIs operating in these objects is not clear. Our comparison with Dong et al. (2015) is not like-for-like, as near-infrared scattered light observations only trace the surface layers of discs. However, Dipierro et al. (2015) have shown that GIs can create strong enough surface density perturbations that could be detected in near-infrared scattered light, hence we feel our agreement with Dong et al. (2015) is noteworthy.

The results of the aforementioned studies, and the results we present in this paper, disagree somewhat with the results of Hall et al. (2016), who use a 1D analytical model (Clarke 2009) extended to 2D and 3D, and find that the spiral structure is detectable only across a narrow parameter space; specifically only at 680 GHz when adopting a distance of 140 pc. Hall et al. (2016) use a cooling time and α that vary locally so they argue that the relative strengths of

the perturbations in their model are much less in the outer part of the disc than they would be in an SPH simulation where β_{cool} is fixed at some relatively low value, such as Dipierro et al. (2014). However, this same argument does not hold for Dong et al. (2015) or our work as non-local global transport is assumed, and hence we argue that GIs should be easier to detect than Hall et al. (2016) postulate.

The aforementioned studies in the literature roughly follow the nomenclature that $m \approx 1/q$ (see Cossins et al. 2009, fig. 8), where m is the number of spiral arms and q is the disc-to-star mass ratio. Moreover, Dong et al. (2015) conclude that m will be a good diagnostic of the total disc mass. Our disc model does not adhere to this approximation, however, as $m = 2$ whilst $q \approx 0.2$, but we emphasize that our disc model is entirely consistent with the long-term behaviour of self-regulated, unstable, radiative discs, particularly considering the highly non-linear behaviour of the system (see the mode analyses in Mejía et al. 2005 and Boley & Durisen 2006, for example, or discussions in Durisen et al. 2007). Therefore, whilst $m \approx 1/q$ can be a useful diagnostic, we note that it is not part of a fundamental criterion for gravitationally unstable discs.

3.3 Observational mass estimates

Disc masses are typically estimated from continuum flux detections using

$$M_{\text{disc}} = \frac{g S_{\nu} d^2}{\kappa_{\nu} B_{\nu}(T)}, \quad (2)$$

where g is the gas-dust mass ratio, usually assumed to be the interstellar medium value of 100, S_{ν} is the flux density, d is the distance to the source, κ_{ν} is an assumed dust opacity and $B_{\nu}(T_{\nu})$ is the Planck function for an assumed dust temperature (Hildebrand 1983). This method assumes that the temperature is uniform and the emission is optically thin. However, young, embedded systems are expected to be optically thick at the wavelengths typically used; for this reason results in the literature usually quote disc masses as lower limits.

Fig. 11 shows the optical depth along the line of sight through our disc model when observed at the two extreme frequencies. Within this regime of the electromagnetic spectrum the dust opacity increases with frequency, as can be seen in Table 1 and Fig. 10, hence we recover the lowest optical depths at the lowest frequencies.

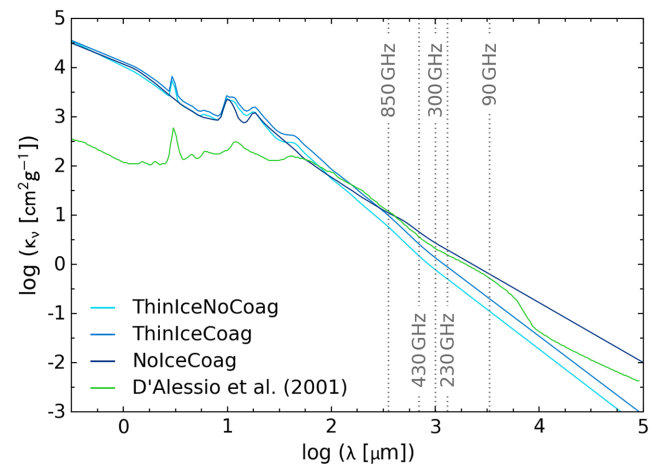


Figure 10. Dust opacity as a function of wavelength for the dust models used to produce synthetic observations of our disc model ('ThinIceNoCoag', 'ThinIceCoag', 'NoIceCoag'), and for the dust model used in the radiative hydrodynamic simulation (D'Alessio et al. 2001).

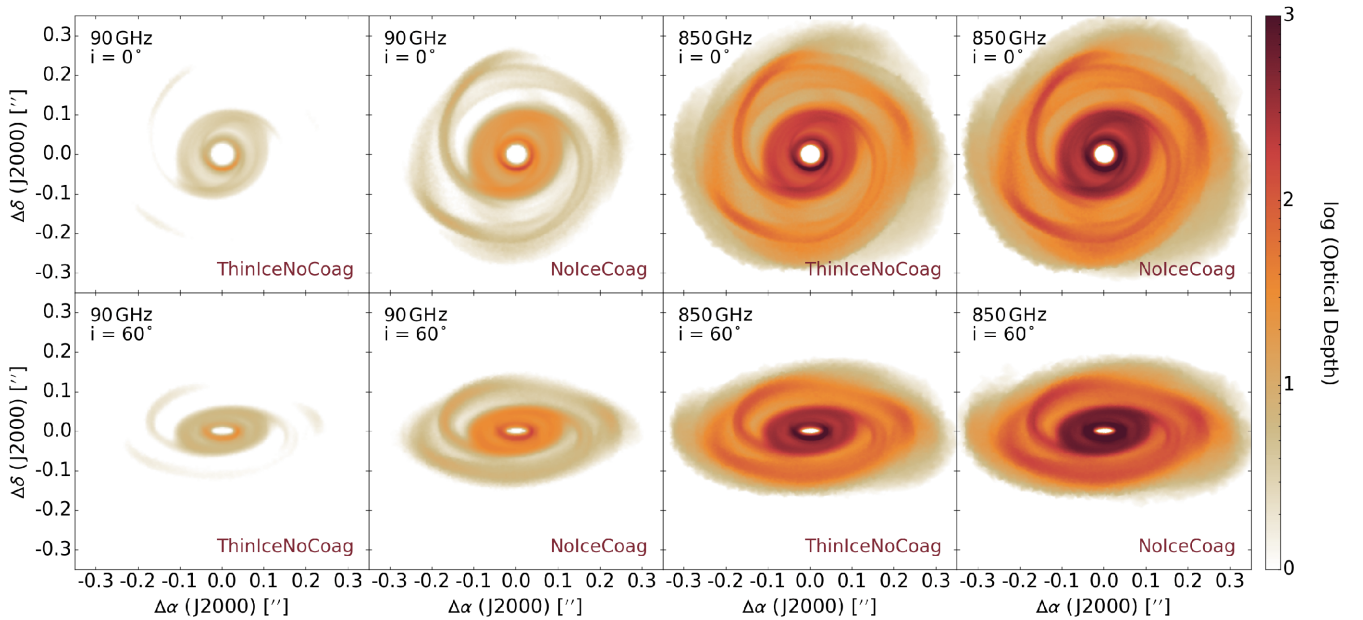


Figure 11. Optical depth along the line of sight of our disc model at two different frequencies, inclinations and dust opacity laws.

Fig. 11 clearly demonstrates that the disc spiral features are optically thick across the majority of our parameter space. To clarify, using the lowest dust opacity configuration (ThinIceNoCoag) at 90 GHz affords an average optical depth of approximately $\tau = 1$ in the face-on spirals, but this increases with inclination, observing frequency and dust opacity. Note that the opacity of the inner disc, which is where most of the disc mass resides, is even higher than in the spirals. Therefore, young, gravitationally unstable discs are likely to be partially optically thick at millimetre wavelengths and substantially optically thick at submillimetre wavelengths, which could result in significantly underestimated disc masses. As we know the exact mass of our disc model, we can explore how inaccurate such disc mass estimates can be.

We assume two different dust temperatures, $T_{\text{dust}} = 20$ and 40 K when calculating the disc mass. The former is chosen to mimic the calculations performed using observations of more evolved, Class II (e.g. Ansdell et al. 2016), based on the median temperature of Class II discs in Taurus-Auriga (Andrews & Williams 2005). However, as GIs can heat young discs significantly (see Paper I), we also adopt a higher dust temperature, $T_{\text{dust}} = 40$ K. Coincidentally, this is the dust temperature Tobin et al. (2016) use when deriving the mass of the recently formed fragment in L1448 IRS3B, as they argue the region around the fragment should be warmed by turbulence. Moreover, this is the peak brightness temperature, T_{B} , towards this fragment, and in optically thick regions $T_{\text{B}} \approx T_{\text{dust}}$.

We estimate the flux density from the Cycle 5 ALMA synthetic observations, shown in Figs A1–A4, by fitting regions to the disc emission that varied with inclination and frequency. We then calculate disc masses across the entire parameter space for our observations produced using both the Cycle 5 antenna configuration and the fully extended, maximally operational antenna configuration. However, as there are minimal differences between the masses derived from the two antenna setups, we omit the fully operational ALMA results. Figs 12 and 13 show how the ratio of the observationally derived mass, M_{obs} , to the actual mass of our simulated disc model, M_{sim} , varies as a function of frequency, inclination, dust opacity and dust temperature. As the blackbody flux, B_{ν} , increases with temperature, we recover a lower disc mass at a higher dust temperature.

We also recover a lower disc mass if we use a dust grain configuration that results in larger optical depths (see Table 1), as we are observing emission from a smaller fraction of the disc. Similarly, if we observe at higher frequencies then we are not able to peer as far through the disc, as can be seen in Fig. 11, so we also recover a lower disc mass. Finally, as the inclination of the disc increases, the recovered mass decreases because the morphology of the disc loses distinction and the line-of-sight column density increases. As a result, the recovered flux density, S_{ν} , decreases, particularly as the orientation approaches edge-on.

In summary of these results, the derived disc mass is sensitive to the disc properties as well as the observational setup. However, some combinations of parameters are probably unrealistic for young, gravitationally unstable discs. For instance, when adopting low dust opacities (ThinIceNoCoag), which assumes no coagulation has occurred at such early stages of disc evolution, and a low dust temperature, $T_{\text{dust}} = 20$ K, the derived disc mass can be significantly larger than the actual mass at low frequencies. Indeed, observing our disc model close to face-on at 90 GHz results in an observed mass, $M_{\text{obs}} = 4M_{\text{sim}}$ when adopting these parameter values. These overestimates occur for two primary reasons. First, using an underestimated dust temperature results in a lower Planck flux when far from the Rayleigh–Jeans limit, which increases the disc mass. As aforementioned, $T_{\text{dust}} = 20$ K is the median temperature of Class II discs and therefore, due to the shock-heating of GIs, is likely an underestimate for young, gravitationally unstable discs. Secondly, this dust configuration (ThinIceNoCoag) disagrees with the underlying dust distribution in the hydrodynamic model across our observational frequency range (see Fig. 10), which assumes dust grains up to 1 mm are present. As observations indicate that larger grains are already present in the envelopes of Class I objects (e.g. Miotello et al. 2014), and are most likely explained by grain growth within the disc (Wong, Hirashita & Li 2016), low dust opacities are likely inaccurate in young, gravitationally unstable discs. Note that whilst assuming grain growth perhaps contradicts our assumption of well mixed dust and gas, we assume there has not been sufficient time for the dust and gas to become decoupled because our disc model is very turbulent.

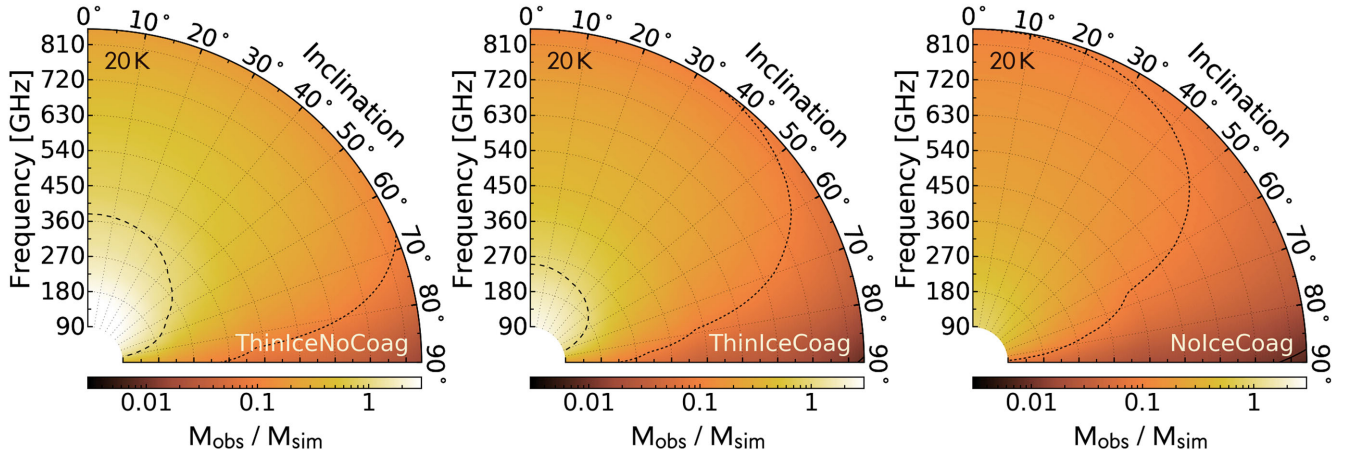


Figure 12. Fraction of actual disc mass derived from observations at various inclinations and frequencies, for three different dust grain properties (see Table 1), assuming a constant dust temperature of 20 K. The contours denote fractions of 0.01 (solid line), 0.1 and 1.0 (longest dashed line). Note that the face-on disc is inclined at 0°.

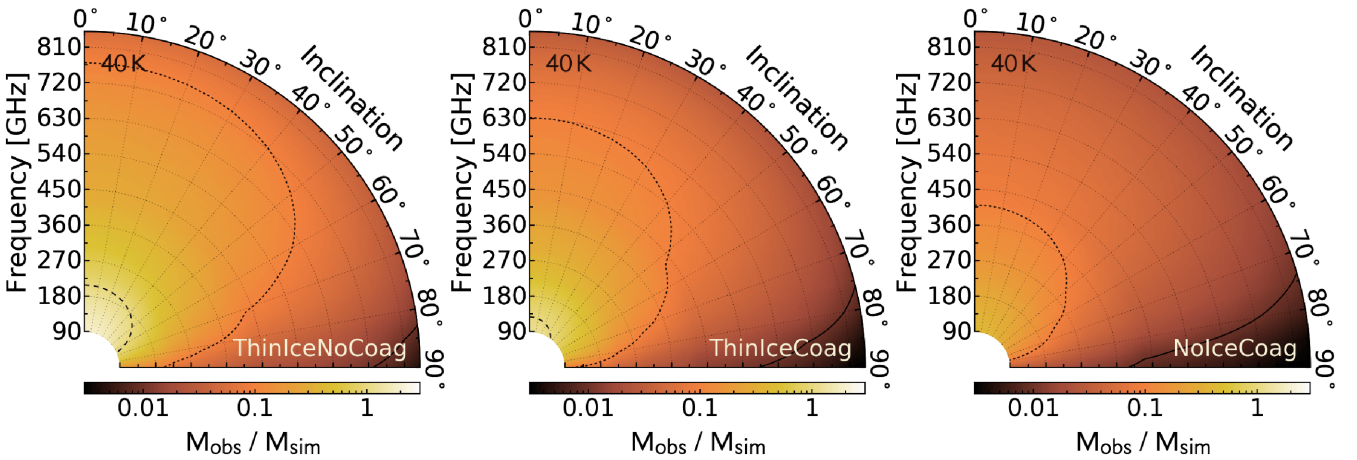


Figure 13. Same as Fig. 12 but assuming a constant dust temperature of 40 K.

Now, as we have argued that a higher dust opacity and dust temperature are most appropriate for our disc model, and we have shown that ‘NoIceCoag’ is self-consistent with the disc simulation, we focus on the right-hand panel of Fig. 13. We recover maximum disc masses when considering a face-on disc, which results in: $M_{\text{obs}} = 0.41 M_{\text{sim}}$ at 90 GHz; $M_{\text{obs}} = 0.21 M_{\text{sim}}$ at 230 GHz; $M_{\text{obs}} = 0.15 M_{\text{sim}}$ at 300 GHz; $M_{\text{obs}} = 0.09 M_{\text{sim}}$ at 430 GHz; and $M_{\text{obs}} = 0.03 M_{\text{sim}}$ at 850 GHz. Deriving disc masses is therefore preferable at lower frequencies. However, as angular resolution is inversely proportional to frequency, obtaining accurate masses and spatially resolving spiral structure are likely mutually exclusive objectives currently. Overall, when using what we understand to be realistic parameters, protoplanetary disc masses derived from continuum observations in the millimetre regime could underestimate the actual mass in a face-on disc by a factor of 2.5–30, depending on observing frequency. Moreover, this underestimate worsens as the disc inclination increases.

3.3.1 Comparison with other studies

Our mass underestimation results are consistent with the predictions of Forgan & Rice (2013), who use self-gravitating disc models to match observations of the Class 0 protostar L1527 IRS, and contin-

uum results of Douglas et al. (2013), who simulate a GI-driven disc representing a violent phase of evolution, with a disc-to-star mass ratio nearly double what we consider. This result also appears consistent with the results of Dunham, Vorobyov & Arce (2014), who report a mass underestimate in a gravitationally unstable disc similar to our model, of a factor of 2–3 at millimetre wavelengths and a factor of 10 or more at submillimetre. Dunham et al. (2014) calculate disc masses assuming a source distance of 250 pc, and so, as we use a source distance of 145 pc, we should perhaps expect less extreme underestimates for our disc model based on their results. However, this is not the case and one possible explanation for this is that the radiative transfer calculations performed by Dunham et al. (2014) are 2D and do not include the full, non-axisymmetric structure of the disc. As a result, spiral features, which we find to be extremely optically thick at most wavelengths (see Fig. 11) are not represented accurately. Despite this, we draw the same conclusions as Dunham et al. (2014) that masses derived from observations of young, embedded discs could be significantly underestimated, except we find that these underestimates may be even more pronounced than they predicted. If this is the case it could have important effects on the derived properties and classifications of young systems. In particular, discs that have been observed and determined not to contain enough mass to be gravitationally unstable could actually be GI-driven.

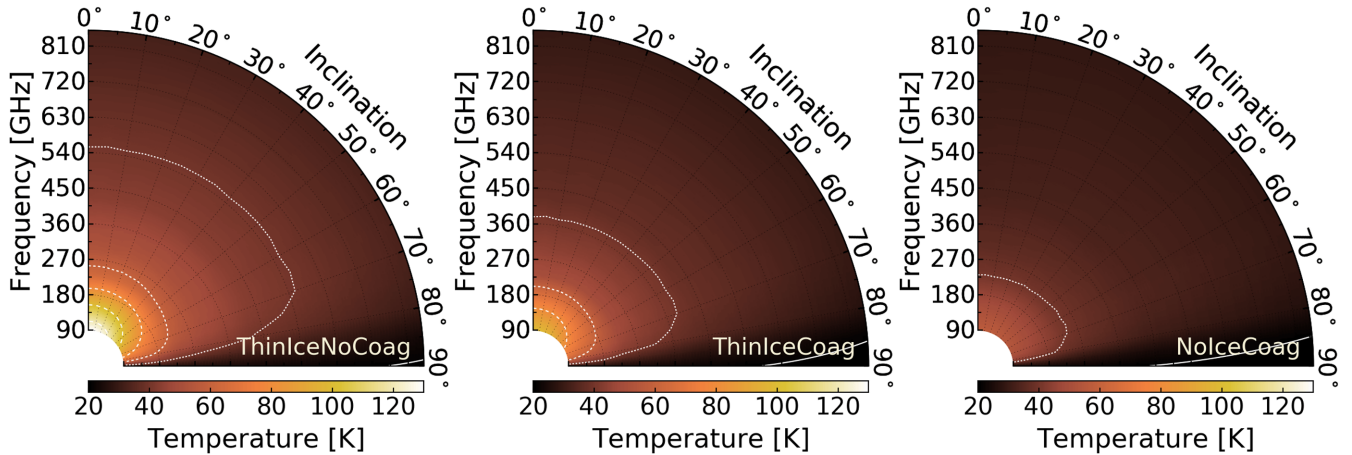


Figure 14. Mass-weighted average temperature of the optically thin emission at various inclinations and frequencies, for three different dust grain properties (see Table 1). The contours denote temperatures from 20 (solid line) to 120 K (longest dashed line) in 20 K increments.

3.3.2 Accurate dust temperatures

It is important to acknowledge that the assumed dust temperature in equation (2) is an estimated average of the optically thin emission temperature, which is typically not well constrained in young, embedded systems. Therefore, it is extremely worthwhile to assess the validity of the dust temperatures we have used in deriving the mass of our disc model thus far, and compare with typical assumptions made when deriving masses from observations.

The actual dust temperature of the optically thin emission varies as a function of the optical depth. Hence, the appropriate average dust temperature to assume when using equation (2) will change as a function of frequency, inclination and dust opacity. We explore this effect in our disc model by calculating the mass-weighted average temperature of the emission above the $\tau = 1$ surface, which characterises the optically thin emission, in a similar method as described in Section 2.5. Fig. 14 demonstrates these results for our disc model and, as can be seen, for all dust opacities we consider, the mass-weighted average temperature at 90 GHz is higher than at shorter wavelengths. This occurs due to the increase in optical depth with frequency, and this effect is much more pronounced when assuming low dust opacities (ThinIceNoCoag) compared to high dust opacities (NoIceCoag). At 270 GHz and beyond, it becomes reasonable to adopt a constant dust temperature, assuming $i < 80^\circ$. At near-edge-on inclinations the morphology of the disc is obscured and the mass-weighted average temperature approaches the minimum value we recover, $T_{\text{dust}} \approx 20$ K.

Beyond Band 3 of ALMA (>116 GHz), our results indicate that the average dust temperature is $T_{\text{dust}} \approx 30$ – 40 K in our GI-driven, embedded protoplanetary disc, except when adopting low dust opacities (ThinIceNoCoag). If observations are performed at Band 3 (84–116 GHz), where the optical depth is lowest, the assumed dust temperature should be increased significantly as more of the shock-heated regions can be observed; for our disc model, $T_{\text{dust}} \approx 55$ – 120 K at 90 GHz, depending on the dust opacity law adopted. Our findings, which are representative of a late Class I/early Class 0 system, disagree with the recommendation of Dunham et al. (2014) to adopt $T_{\text{dust}} \approx 30$ K in a Class 0 source and $T_{\text{dust}} \approx 15$ K in a Class I source. This is most likely because, as aforementioned, their radiative transfer modelling neglects the shock heating driven by GIs, and so their temperature estimates of the optically thin emission

could be underestimated. This would then mean that their derived masses are likely to be overestimated.

Overall, the method for deriving masses from continuum flux observations is questionable for two primary reasons. First, the dust temperatures in a gravitationally unstable disc will cover a large range (approximately 20–250 K in our disc model), and reducing this to a mass-weighted average dust temperature neglects the complex non-axisymmetric structure of the disc. Secondly, the dust temperature is an uncertain property as it is dependent on the dust opacity, which itself is poorly constrained due to the unconfirmed state of grain growth in young, embedded systems. Moreover, using a dust temperature that is inappropriately low for the system in question can inadvertently compensate for the large optical depths caused by large dust opacities. For example, we can recover $M_{\text{obs}} \approx M_{\text{sim}}$ for our disc model when adopting large opacities (NoIceCoag) if we use $T_{\text{dust}} = 20$ K, which as we have shown is an unrealistic dust temperature for our disc model. Therefore, even in this case, where the observational mass approaches the actual mass, the calculated disc mass is not accurate. Hence, mass estimates derived from continuum flux observations of embedded, gravitationally unstable discs are inherently unreliable.

4 CONCLUSIONS

4.1 LIME optimizations

We have produced continuum emission maps of a protoplanetary disc model using LIME and in Section 2 have explored in-depth how accurate images can be attained. The main results from this body of work, which focuses on continuum images, are as follows:

(i) The optimal grid point weighting function is dependent on the model used and features of interest within the model. We use number density weighting, which is the vanilla LIME default, and find that $n_0 = 1 \times 10^{12} \text{ cm}^{-3}$ and $w = 0.5$ highlight the spiral features within our disc model.

(ii) A convergence test should be used to find the number of grid points necessary to produce an accurate continuum emission image. This will be dependent on the model, but for our young, gravitationally unstable protoplanetary disc model, we find that a minimum of 2×10^5 grid points should be used. Note, however,

that more grid points are always preferable if the computational cost can be afforded. This strategy should also be applied to line images.

(iii) The contribution to emission intensity in high optical regions is negligible. Grid points can therefore be omitted where $\tau > 3$, which, as these are the densest regions, increases the number of grid points in the observed regions significantly. As a result, this method results in an improvement in the accuracy of the continuum images for our disc model when compared to vanilla LIME images produced using the same number of grid points. Hence, these findings demonstrate that specialized grid construction can improve the efficiency and accuracy of LIME images.

(iv) Averaging multiple LIME runs using identical inputs improves the accuracy of the continuum emission image because the grid point coverage is effectively increased. As multiple instances of LIME can be run simultaneously, we see no strong argument against averaging consistent LIME images, barring strict computational or time restraints.

Although LIME is specialized for line emission, we have presented our findings as an indicator that care should be taken even when using a radiative transfer code to produce continuum emission images, in order to ensure these images are accurate. When considering line emission images, the issues we have discussed in this paper are amplified due to the complexity of the radiative transfer equation for molecular line transitions. Therefore, even more care should be taken in this scenario. Our results regarding the continuum should be applicable to molecular line images, which we will explore in a forthcoming publication.

4.2 Continuum emission and mass estimates

In Section 2, we produced synthetic observations of our disc model using CASA, across a large parameter space consisting of different frequencies, inclinations and dust opacities. Our model represents a $0.17 M_{\odot}$ protoplanetary disc that surrounds a $0.8 M_{\odot}$ protostar likely to evolve into a Solar-like star. As a result our work may be indicative of observations of an object similar to our early Solar system. The main conclusions we have drawn from these synthetic observations are as follows:

(i) Using the antenna configuration for ALMA Cycle 5 and observing for 60 min reveals that, at 90–300 GHz, spiral features can be readily detected in a 0° – 60° inclined young, gravitationally unstable disc at a distance of 145 pc. Spiral features can also be distinguished at near-face-on discs at 430 GHz, but can only be identified at 850 GHz when the dust opacity is very low. For Cycle 5, we find that the optimal frequency to observe non-axisymmetric structure in embedded objects is 230 GHz (1.3 mm).

(ii) Using a fully extended, maximally operational ALMA affords similar results to Cycle 5 observations at lower frequencies as the angular resolution and sensitivity do not change significantly. However, at 430 and 850 GHz the angular resolution is much improved; we obtain a minimum spatial resolution of 1.6 and 0.8 au at 430 and 850 GHz, respectively. Unfortunately, the spirals are still not distinguishable at these higher angular resolutions because the noise level increases with frequency. Nevertheless, the non-axisymmetric inner disc structure can be seen in extraordinary detail, particularly in inclined discs. Hence, once ALMA reaches maximum capacity, observations in Bands 8, 9 and 10 could be used to probe the innermost regions of embedded discs and potentially offer unprecedented insight into their evolution mechanisms.

(iii) We find that the majority of our gravitationally unstable disc model is substantially optically thick at frequencies beyond

90 GHz. This is because the spiral waves and inner disc contain a considerable amount of mass in vertically elongated regions, which results in large line-of-sight optical depths. As a result, observations of embedded discs should be performed at long wavelengths where the optical depth is lowest.

(iv) As the GIs within a gravitationally unstable disc heat the disc material on a global scale, using a dust temperature of 20 K, which is typically used for more evolved discs, is likely inaccurate. Instead, we find that the mass-weighted average temperature of the optically thin regions within our disc model is $T_{\text{dust}} \approx 30$ – 40 K at high frequencies (≥ 230 GHz). At low frequencies (90 GHz), $T_{\text{dust}} \approx 55$ – 90 K as the disc is less optically thick, which means more of the shock-heated spiral structure is observed.

(v) Using the flux density of observations to estimate the mass of young, embedded discs is an inherently flawed procedure for two primary reasons. First, condensing the 3D spiral structure into one value for the dust temperature and one value for the dust opacity across the entire disc is misrepresentative. Secondly, the dust temperature and dust opacity in optically thick objects are largely unknown quantities, hence mass derivations are unreliable.

(vi) Assuming what is believed to be an appropriate dust temperature and opacity within a young, gravitationally unstable disc, the disc mass derived from observations can be underestimated by a factor of 2.5–30 in a face-on disc, depending on frequency. Moreover, this underestimate worsens as the disc inclination increases. If our assumptions are appropriate, then this could retrospectively validate GIs in discs previously thought not massive enough to be gravitationally unstable, which could have a significant impact on the understanding of the formation and evolution of protoplanetary discs.

The results of our synthetic observations are incredibly encouraging for the future and suggest that spiral features in young, embedded, gravitationally unstable discs should be detectable by recent and future cycles of ALMA at millimetre wavelengths, and perhaps at submillimetre wavelengths given optimal conditions and long observation times. Our synthetic observations support the recent detection of spiral structure in protoplanetary discs (e.g. Pérez et al. 2016; Tobin et al. 2016), which strengthens the argument that these non-axisymmetric features are driven by GIs. However, as these objects are more complex than our disc model, this argument is still tentative. Therefore, more observations are needed to assess the validity of GIs in young, embedded objects. If such observations are performed across multiple wavelengths, we should be able to attain unprecedented levels of detail of the disc morphology, which will allow the importance of GIs on the formation and evolution of protoplanetary discs to be understood.

The synthetic observations of our model disc suggest that the mass of a young, gravitationally unstable disc cannot be accurately constrained as these objects are likely optically thick across a broad range of ALMA frequencies. Therefore alternative approaches should be adopted, such as extracting kinematic information from molecular line observations. Moreover, when considering line emission, observing spiral structure in certain, optically thin, molecular tracers could suggest that localized heating of the disc material is occurring (see Paper I), which would strengthen the argument for GI-driven spiral structure. Therefore, in the near future, we will use LIME and our gridding optimizations to explore the ability of ALMA to extract kinematic information and resolve spiral structure in commonly observed molecular line transitions, in order to quantify the accuracy of disc masses derived from line images and to support the findings of our synthetic continuum observations.

ACKNOWLEDGEMENTS

We would like to thank the anonymous referee for constructive comments that have improved the clarity of this manuscript. MGE gratefully acknowledges a studentship from the European Research Council (ERC; project PALs 320620). JDI gratefully acknowledges support from the DISCSIM project, grant agreement 341137, funded by the European Research Council under ERC-2013-ADG. TWH, PC and LSz acknowledge the financial support of the European Research Council (ERC; project PALs 320620). ACB's contribution was supported, in part, by The University of British Columbia and the Canada Research Chairs program.

REFERENCES

- Andrews S. M., Williams J. P., 2005, *ApJ*, 631, 1134
 Andrews S. M., Rosenfeld K. A., Kraus A. L., Wilner D. J., 2013, *ApJ*, 771, 129
 Ansdell M. et al., 2016, *ApJ*, 828, 46
 Bae J., Hartmann L., Zhu Z., Nelson R. P., 2014, *ApJ*, 795, 61
 Benisty M. et al., 2015, *A&A*, 578, L6
 Boley A. C., Durisen R. H., 2006, *ApJ*, 641, 534
 Boley A. C., Durisen R. H., 2008, *ApJ*, 685, 1193
 Boley A. C., Durisen R. H., Nordlund Å., Lord J., 2007, *ApJ*, 665, 1254
 Brinch C., Hogerheijde M. R., 2010, *A&A*, 523, A25
 Cai K., Durisen R. H., Boley A. C., Pickett M. K., Mejía A. C., 2008, *ApJ*, 673, 1138
 Clarke C. J., 2009, *MNRAS*, 396, 1066
 Cossins P., Lodato G., Clarke C. J., 2009, *MNRAS*, 393, 1157
 Cossins P., Lodato G., Testi L., 2010, *MNRAS*, 407, 181
 D'Alessio P., Calvet N., Hartmann L., 2001, *ApJ*, 553, 321
 D'Alessio P., Calvet N., Hartmann L., Franco-Hernández R., Servín H., 2006, *ApJ*, 638, 314
 Dipierro G., Lodato G., Testi L., de Gregorio Monsalvo I., 2014, *MNRAS*, 444, 1919
 Dipierro G., Pinilla P., Lodato G., Testi L., 2015, *MNRAS*, 451, 974
 Dong R., Hall C., Rice K., Chiang E., 2015, *ApJ*, 812, L32
 Douglas T. A., Caselli P., Ilee J. D., Boley A. C., Hartquist T. W., Durisen R. H., Rawlings J. M. C., 2013, *MNRAS*, 433, 2064
 Dullemond C. P., Juhasz A., Pohl A., Sereshti F., Shetty R., Peters T., Commercon B., Flock M., 2012, *Astrophysics Source Code Library*, record ascl:1202.015
 Dunham M. M., Vorobyov E. I., Arce H. G., 2014, *MNRAS*, 444, 887
 Durisen R. H., Boss A. P., Mayer L., Nelson A. F., Quinn T., Rice W. K. M., 2007, in Reipurth B., Jewitt D., Keil K., eds, *Protostars and Planets V*, Univ. Arizona Press, Tucson, p. 607
 Evans M. G., Ilee J. D., Boley A. C., Caselli P., Durisen R. H., Hartquist T. W., Rawlings J. M. C., 2015, *MNRAS*, 453, 1147 (Paper I)
 Forgan D., Rice K., 2013, *MNRAS*, 433, 1796
 Forgan D., Rice K., Stamatellos D., Whitworth A., 2009, *MNRAS*, 394, 882
 Greaves J. S., Rice W. K. M., 2011, *MNRAS*, 412, L88
 Hall C., Forgan D., Rice K., Harries T. J., Klaassen P. D., Biller B., 2016, *MNRAS*, 458, 306
 Harker D. E., Desch S. J., 2002, *ApJ*, 565, L109
 Harsono D., Alexander R. D., Levin Y., 2011, *MNRAS*, 413, 423
 Hildebrand R. H., 1983, *QJRAS*, 24, 267
 Ilee J. D., Boley A. C., Caselli P., Durisen R. H., Hartquist T. W., Rawlings J. M. C., 2011, *MNRAS*, 417, 2950
 Keto E., Rawlings J., Caselli P., 2014, *MNRAS*, 440, 2616
 Kratter K., Lodato G., 2016, *ARA&A*, 54, 271
 Kratter K. M., Murray-Clay R. A., 2011, *ApJ*, 740, 1
 Lodato G., Rice W. K. M., 2004, *MNRAS*, 351, 630
 Machida M. N., Matsumoto T., 2011, *MNRAS*, 413, 2767
 McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, *Astronomical Society of the Pacific Conference Series Vol. 376, Astronomical Data Analysis Software and Systems XVI*, p. 127
 Mejía A. C., Durisen R. H., Pickett M. K., Cai K., 2005, *ApJ*, 619, 1098
 Meru F., Juhász A., Ilee J. D., Clarke C. J., Rosotti G. P., Booth R. A., 2017, *ApJ*, 839, L24
 Miotello A., Testi L., Lodato G., Ricci L., Rosotti G., Brooks K., Maury A., Natta A., 2014, *A&A*, 567, A32
 Muto T. et al., 2012, *ApJ*, 748, L22
 Ossenkopf V., Henning T., 1994, *A&A*, 291, 943
 Pérez L. M. et al., 2012, *ApJ*, 760, L17
 Pérez L. M. et al., 2016, *Science*, 353, 1519
 Richert A. J. W., Lyra W., Boley A., Mac Low M.-M., Turner N., 2015, *ApJ*, 804, 95
 Tobin J. J. et al., 2016, *Nature*, 538, 483
 Tomida K., Machida M. N., Hosokawa T., Sakurai Y., Lin C. H., 2017, *ApJ*, 835, L11
 Wong Y. H. V., Hirashita H., Li Z.-Y., 2016, *PASJ*, 68, 67

APPENDIX A: CYCLE 5 CONTINUUM EMISSION IMAGES AT 90, 300, 430 AND 850 GHz

Figs A1–A4 represent the capabilities of ALMA Cycle 5 when using our disc model and observational setup, with an integration time of 1 h. We use the antenna configuration that is available and affords us the best angular resolution at each frequency, i.e. 90 GHz: C43-10, 300 GHz: C43-8, 430 GHz: C43-7 and 850 GHz: C43-7. As can be seen, the spiral structure is distinguishable in a face-on disc across all frequencies and dust opacities. At an inclination of 60° the spirals are only identifiable at 300 GHz or lower, and only convincingly when assuming a lower dust opacity.

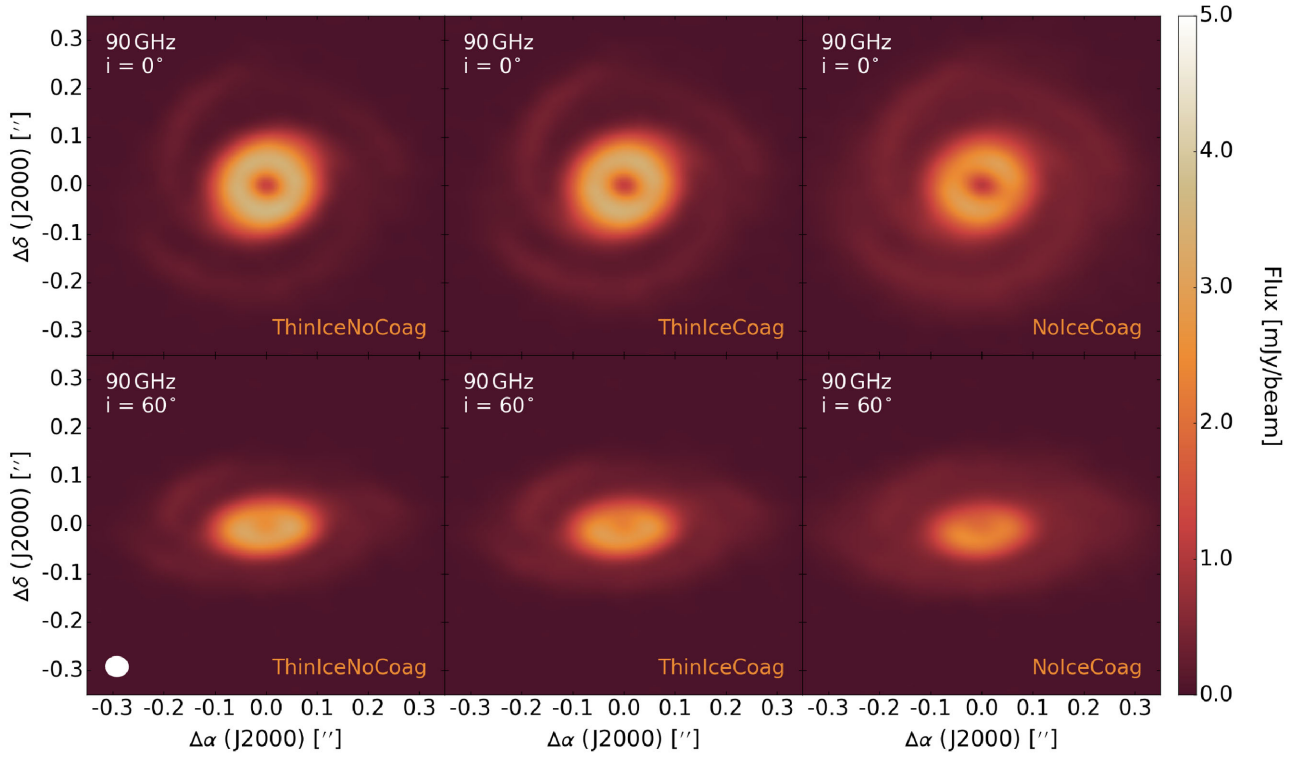


Figure A1. Continuum emission images of our disc model at 90 GHz for three different dust configurations and two different inclinations, synthesized using the ALMA Cycle 5 antenna configuration. The white ellipse in the lower left indicates the size of the beam, which is 0.046×0.044 arcsec² and is constant across all panels. The noise is approximately $11 \mu\text{Jy beam}^{-1}$ across the parameter space.

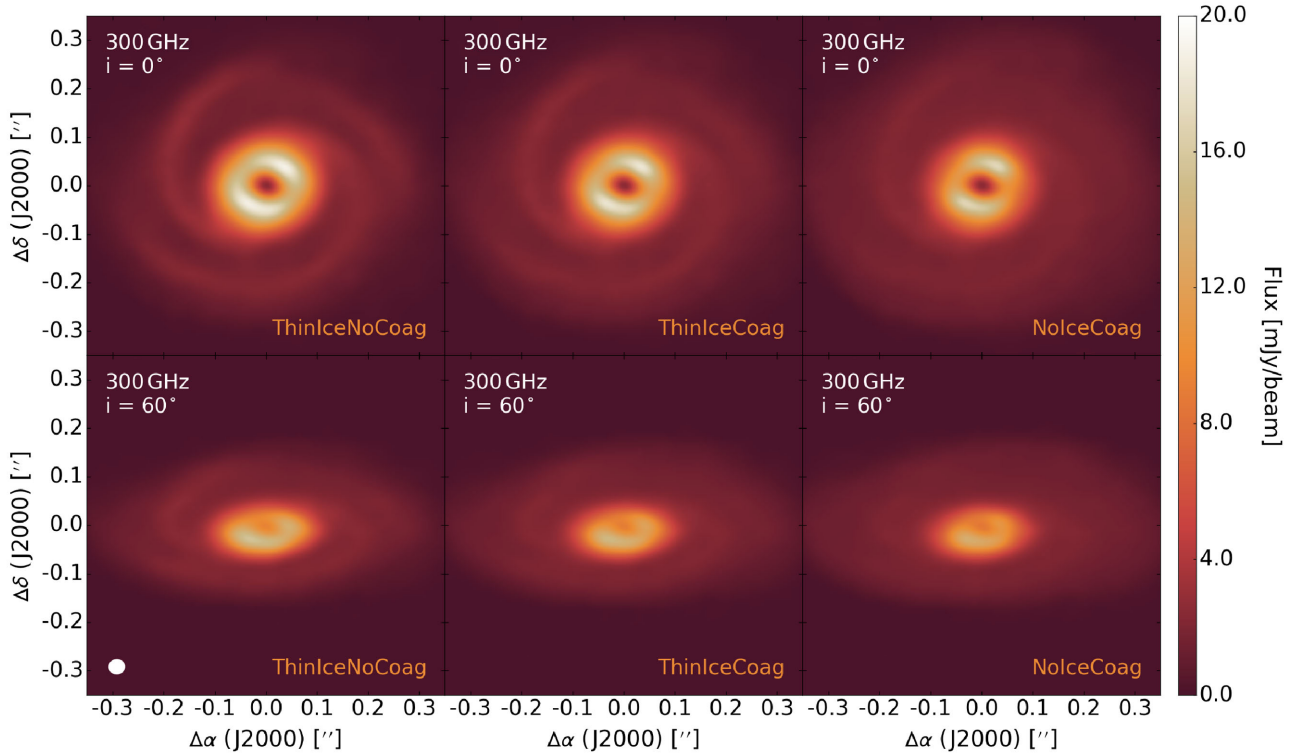


Figure A2. Same as Fig. A1 but synthesized at 300 GHz with a beamsize of 0.033×0.031 arcsec² and a noise level of approximately $26 \mu\text{Jy beam}^{-1}$.

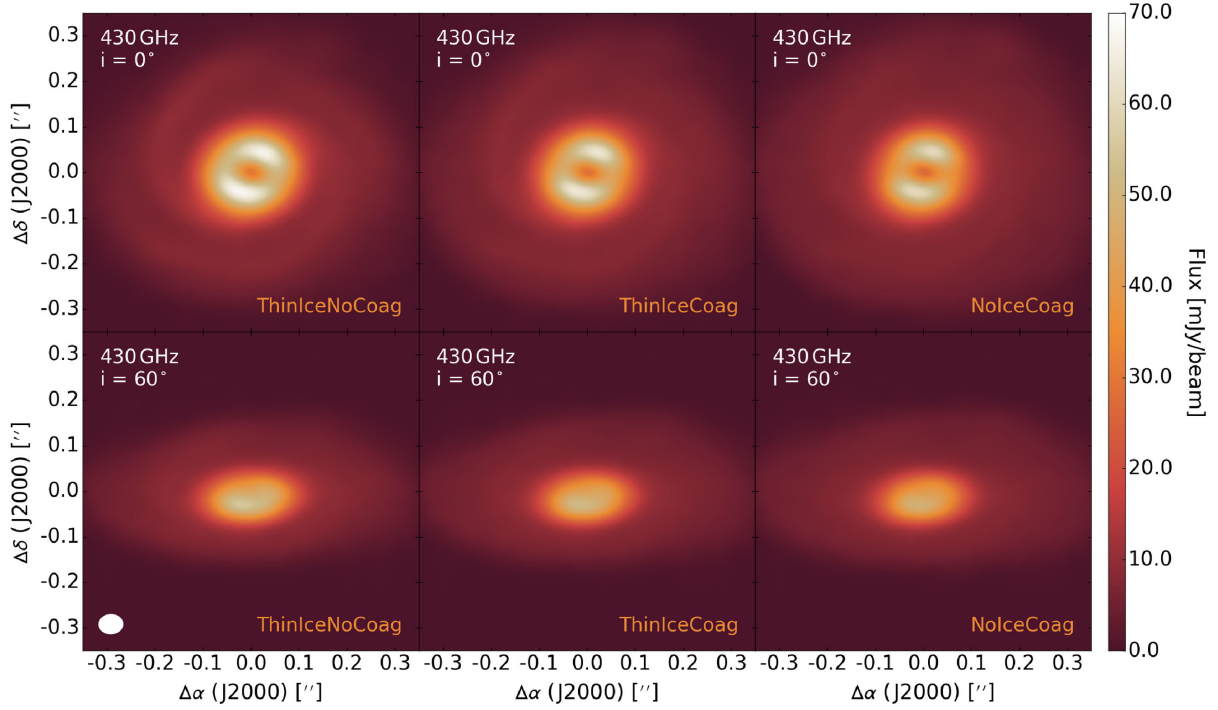


Figure A3. Same as Fig. A1 but synthesized at 430 GHz with a beamsize of 0.052×0.045 arcsec² and a noise level of approximately $115 \mu\text{Jy beam}^{-1}$.

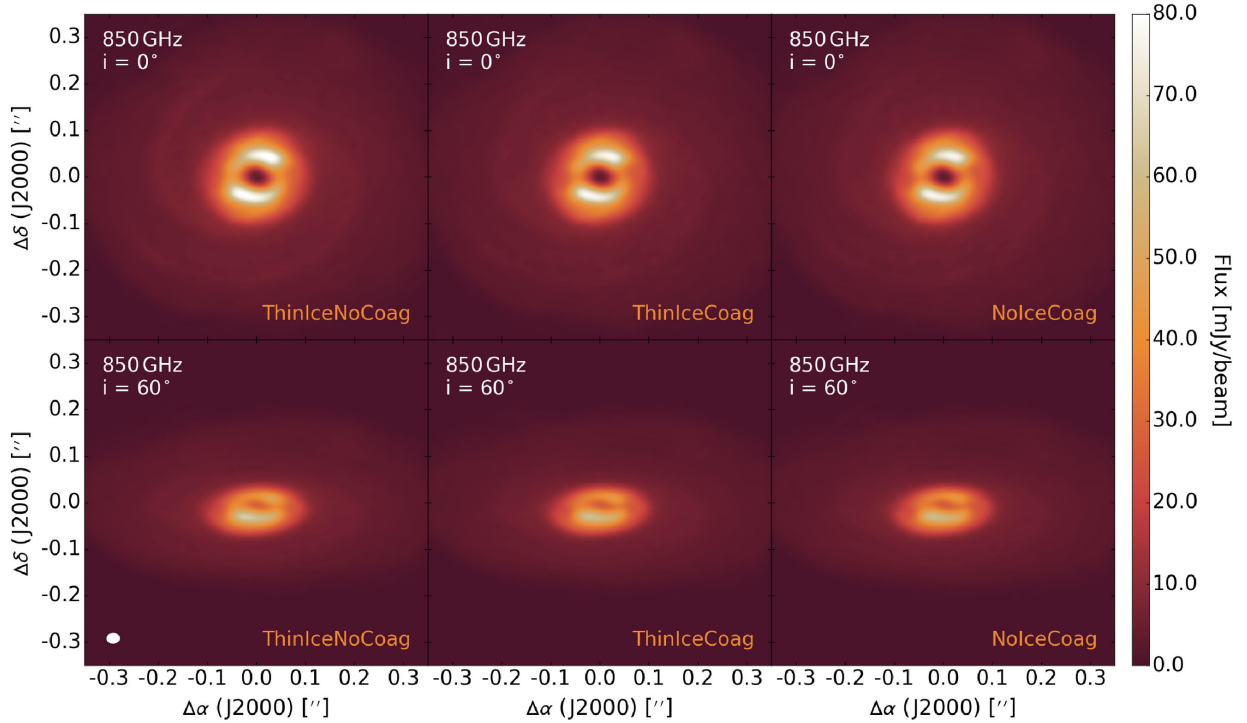


Figure A4. Same as Fig. A1 but synthesized at 850 GHz with a beamsize of 0.027×0.023 arcsec² and a noise level of approximately $110 \mu\text{Jy beam}^{-1}$.

APPENDIX B: FULLY OPERATIONAL ALMA CONTINUUM EMISSION IMAGES AT 300, 430 AND 850 GHz

Figs B1–B3 represent the highest possible image fidelity ALMA could achieve using our observational setup described in Section 3,

assuming the maximum baseline of 16.2 km is available at all frequencies. As can be seen, the spiral structure is very prominent at 300 GHz, with an $S/N > 30$, due to the beamsize reduction at higher frequencies. Unfortunately, the sensitivity is inversely proportional to the frequency, hence at 430 GHz the spirals are not as easily identified because noise becomes significant. In fact, because of

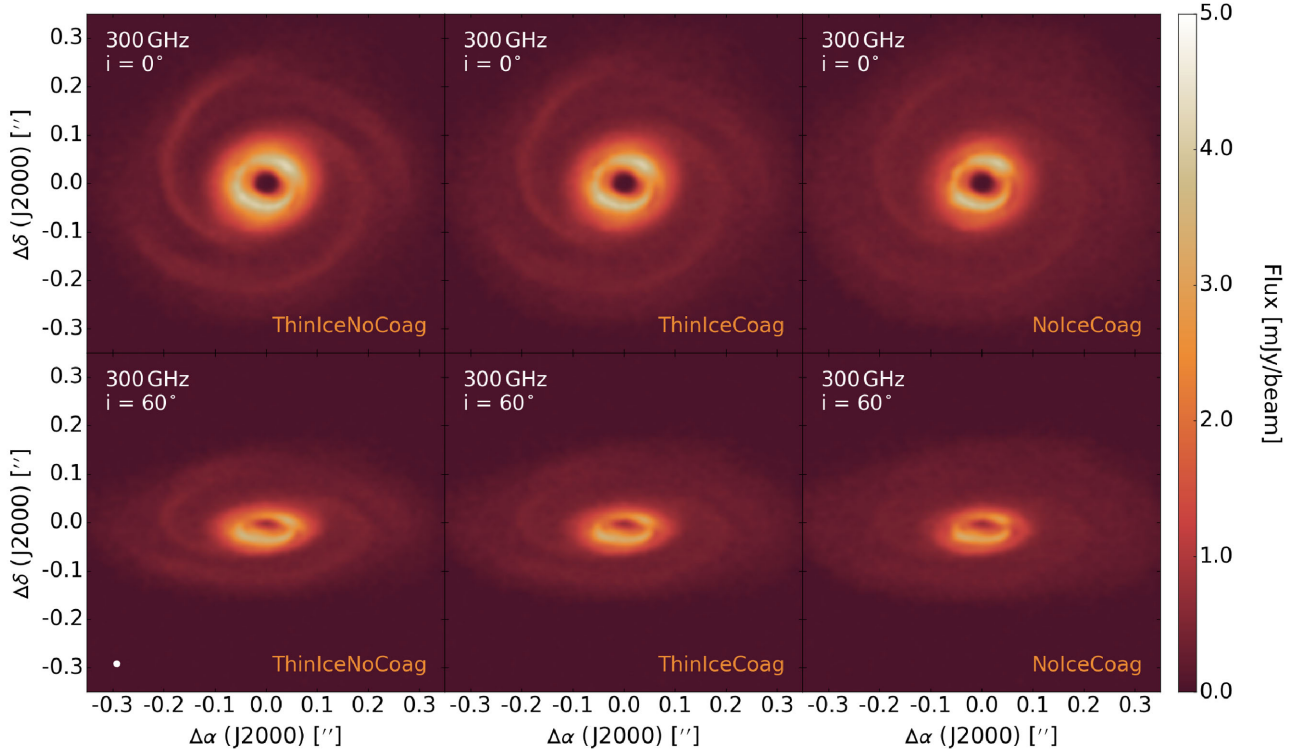


Figure B1. Continuum emission images of our disc model at 300 GHz for three different dust configurations and two different inclinations, synthesized using a fully extended, maximally operational ALMA antenna configuration. The white ellipse in the lower left-hand panel indicates the size of the beam, which is $0.014 \times 0.014 \text{ arcsec}^2$ and is constant across all panels. The noise is approximately $20 \mu\text{Jy beam}^{-1}$ across the parameter space.

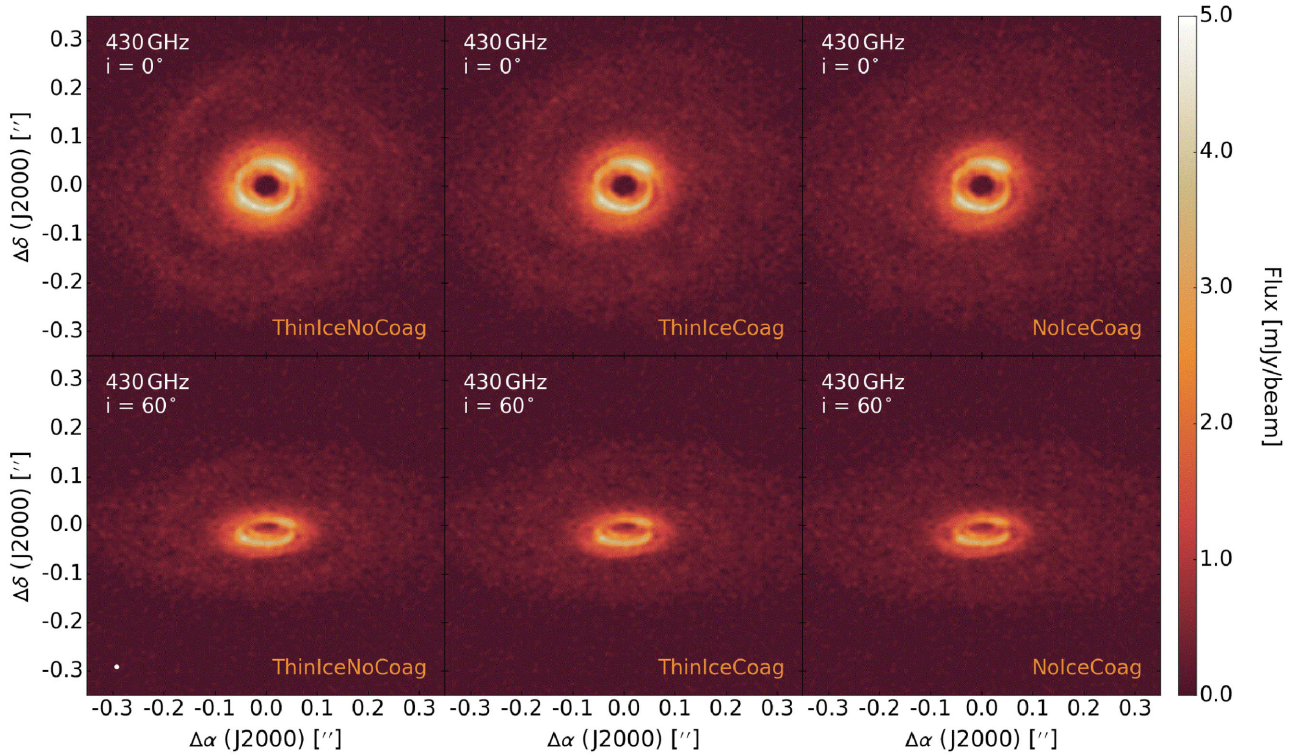


Figure B2. Same as Fig. B1 but synthesized at 430 GHz with a beamsize of $0.010 \times 0.010 \text{ arcsec}^2$ and a noise level of approximately $85 \mu\text{Jy beam}^{-1}$.

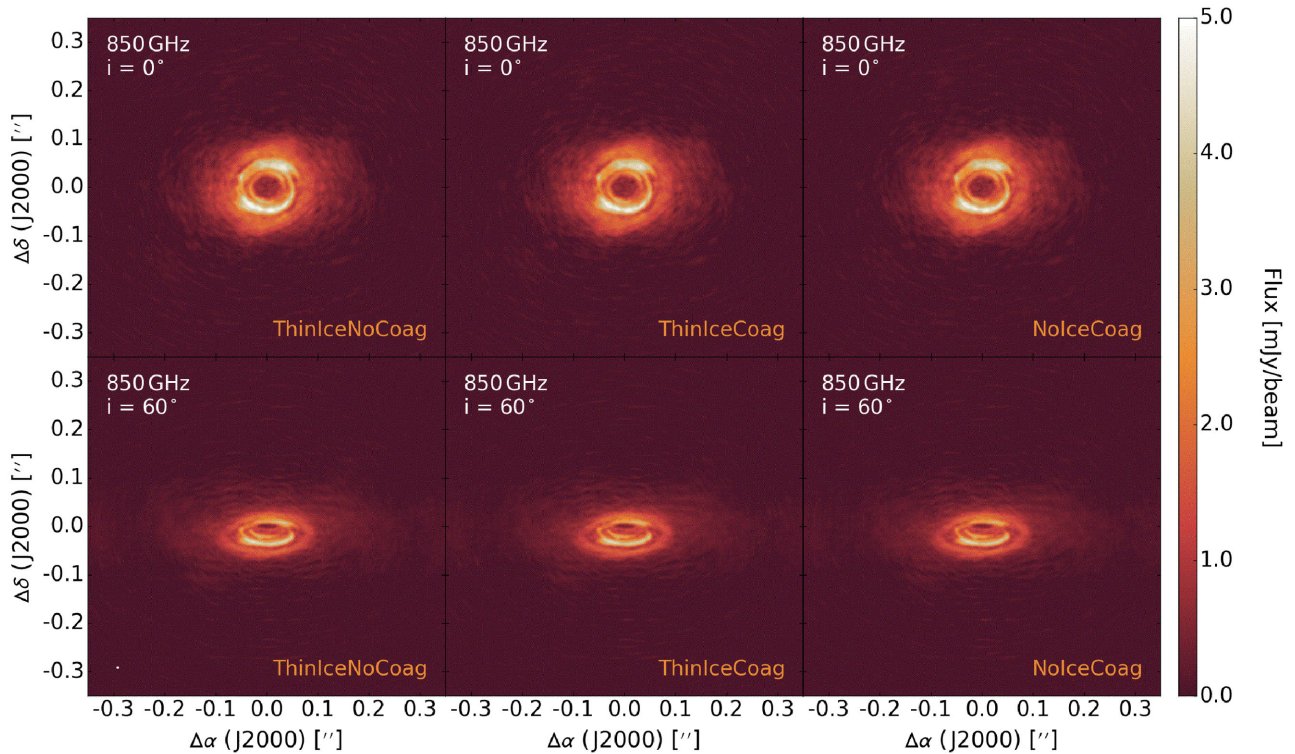


Figure B3. Same as Fig. B1 but synthesized at 850 GHz with a beamsize of 0.005×0.005 arcsec² and a noise level of approximately $25 \mu\text{Jy beam}^{-1}$.

this, the spirals are only distinguishable when assuming a very low dust opacity at 430 GHz, and at 850 GHz the spirals are completely dominated by the noise. Note that we do not show the observations at 90 and 230 GHz because they are virtually identical to those performed using the ALMA Cycle 5 antenna configuration as the maximum baseline is already at the fully operational limit for Bands 3 and 6; the only difference is the number of antennas (42 for Cycle 5, 50 for fully operational ALMA).

APPENDIX C: COMPARISON WITH RADMC-3D

In this work, we have used LIME to produce continuum emission maps that we have then observed synthetically. However, there are numerous radiative transfer codes and it is very important to compare results between them using consistent models in order to understand the strengths and limitations of each. One such alternative to LIME is RADMC-3D, which is a Monte Carlo radiative transfer code used to produce dust continuum emission images and gas line emission images in LTE regions. RADMC-3D can also model non-LTE line emission, but currently, only if local modes are adopted. We input our disc model into RADMC-3D using our pre-computed hydrodynamic dust temperatures, adopt the ‘ThinIceCoag’ dust opacities and produce a continuum flux image at 300 GHz to compare to the LIME image produced from eight averaged runs using our optimal setup ($n_0 = 1 \times 10^{12} \text{ cm}^{-3}$, $w = 0.5$, 2×10^5 points, $z \geq z_{\tau=3}$).

Fig. C1 shows the residual flux between our optimal LIME image and the RADMC-3D image, and indicates that the two codes produce compatible images as there is a residual difference of <10 per cent within the spiral arms. However, there are large discrepancies, up to ≈ 60 per cent, in the inter-arm and innermost disc regions that are most likely due to the difference in gridding employed by LIME and RADMC-3D. In RADMC-3D, the radiative transfer grid is the entire

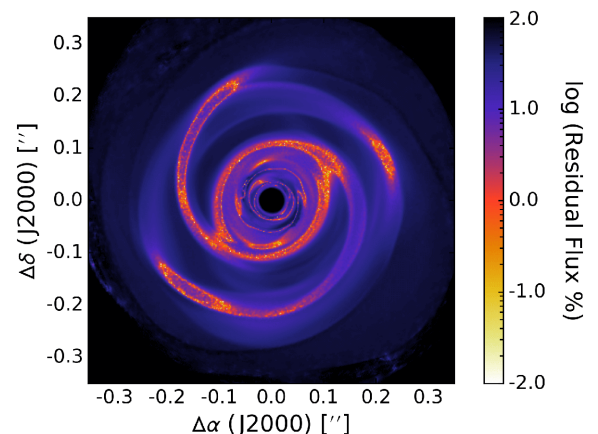


Figure C1. Residual flux between LIME runs using 2×10^6 grid points, f_L , and RADMC-3D, f_R , using the full hydrodynamic grid. We define this residual as $(f_L - f_R)/(f_L + f_R)$, where f denotes flux per pixel.

hydrodynamic model grid, whereas in LIME the points are subsampled from this grid and density weighted. Therefore, in less dense regions, and hence optically thinner regions, the grid point spacing in RADMC-3D is much finer than in our LIME runs. For instance, in the inter-arm regions at the disc mid-plane, $l > 1.0$ au, whereas the RADMC-3D grid resolution is 0.25 au; it is important to understand that the mean free path is a statistical average and hence a smaller grid point separation is always preferable if it is achievable. Conversely, in the innermost disc region, where the optical depth is considerably higher, the grid spacing in LIME becomes finer than that in RADMC-3D. This may explain why both the interarm and innermost disc flux is lower in the RADMC-3D continuum emission map, but for opposite reasons.

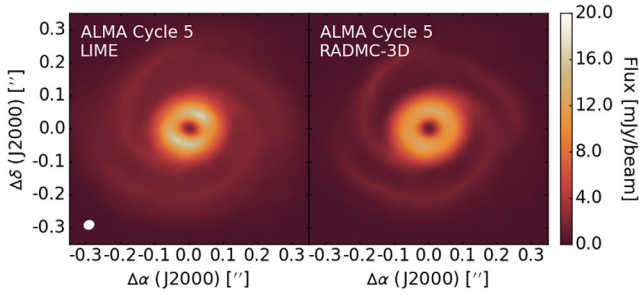


Figure C2. Continuum emission images of LIME (left-hand side) and RADMC-3D (right-hand side) radiative transfer calculations of our disc model at 300 GHz and assuming ‘ThinIceCoag’ dust opacities, synthesized using the ALMA Cycle 5. The white ellipse in the lower left-hand side indicates the size of the beam, which is 0.029 arcsec and is constant across all panels. The noise is approximately $40 \mu\text{Jy beam}^{-1}$ across the parameter space.

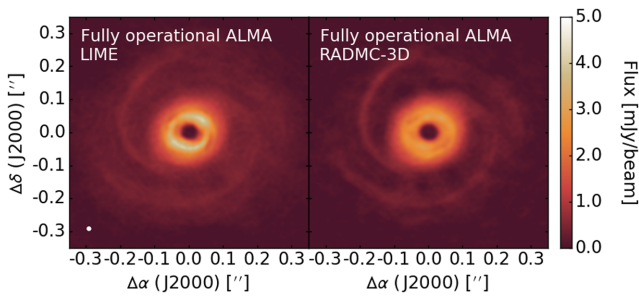


Figure C3. Continuum emission images of LIME (left-hand side) and RADMC-3D (right-hand side) radiative transfer calculations of our disc model at 300 GHz and assuming ‘ThinIceCoag’ dust opacities, synthesized using the ALMA Cycle 5. The white ellipse in the lower left-hand side indicates the size of the beam, which is 0.016 arcsec and is constant across all panels. The noise is approximately $35 \mu\text{Jy beam}^{-1}$ across the parameter space.

In order to determine if this discrepancy in inter-arm and innermost disc flux leads to observational differences, we produce synthetic observations of our LIME and RADMC-3D continuum emission maps, using a face-on inclination and the same observational setup as aforementioned. The results for ALMA Cycle 5 and a fully extended, maximally operational ALMA are shown in Figs C2 and C3,

respectively, and demonstrate that, whilst the spiral fluxes are consistent between the two radiative transfer codes as expected, the innermost disc flux is lower in the RADMC-3D observations. More significantly for our purposes, however, is that the contrast between arm and inter-arm regions is enhanced in the RADMC-3D observations, which results in a more easily detected spiral structure. As this is most likely a result of finer gridding in RADMC-3D, we conclude that one should perhaps use RADMC-3D to produce continuum emission maps when a high-resolution model is available.

Whilst RADMC-3D likely produces more accurate continuum emission maps than LIME for our disc model, at least in the inter-arm regions, we have opted to show results for LIME throughout this paper for a number of reasons. First, the continuum emission maps we have produced have only utilized raytracing as the dust temperatures are known a priori from the radiative hydrodynamic model. As a result, the computational cost of using a high resolution, regularly spaced grid is low. However, if one were to compute the dust temperatures with a Monte Carlo radiative transfer code, then the computational time would increase significantly as enough photon packages need to be propagated to ensure a sufficient number enter each cell, which becomes especially problematic in the optically thick regions. In this case, using our optimal LIME grid may be more computationally efficient than using the high resolution, regularly spaced RADMC-3D grid. Secondly, LIME is one of the best available tools for producing line emission maps because LIME can calculate level populations self-consistently, without the assumption of LTE or optically thin line emission, or the use of the large velocity gradient method that RADMC-3D relies upon. Note, however, that whether these assumptions would result in large errors in emission maps is not clear. Nevertheless, when producing line emission maps, the computational time no longer scales linearly with the number of grid points as the level populations for each grid point must be calculated in an iterative sense until convergence is reached. Consequently, it becomes much more important efficiency-wise to optimize the grid sampling, which can be achieved by, for example, restricting the grid points in relation to the continuum optical depth surface (see Section 2.5). We plan to explore this avenue of research in a future publication.

This paper has been typeset from a \LaTeX file prepared by the author.